AD-A094 899 SCHOOL OF AEROSPACE MEDICINE BROOKS AFB TX F/G 12/1
ON TWO-SIDED CONFIDENCE AND TOLERANCE LIMITS FOR NORMAL DISTRIB--ETC(U)
DEC 80 A J RAHE
UNCLASSIFIED SAM-TR-80-46 NL

1 OF 1
AD
A094899

END
DATE
FILMED
3-81
DTIC

(14) Report SAM-TR-80-46

LEVEL

(12)

AD A094899

(6) ON TWO-SIDED CONFIDENCE AND TOLERANCE LIMITS FOR NORMAL DISTRIBUTIONS

(10) Alton J. Rahe, M.S.

(11) Dec[redacted]980 (12) 81

(9) Final Report, [redacted] Jan[redacted]967— May 1967

(16) 7930

(17) 15

USAF SCHOOL OF AEROSPACE MEDICINE
Aerospace Medical Division (AFSC)
Brooks Air Force Base, Texas 78235

NOTICES

This final report was submitted by personnel of the Advanced Analysis Branch, Data Sciences Division, USAF School of Aerospace Medicine, Aerospace Medical Division, AFSC, Brooks Air Force Base, Texas, under job order 7930-15-02.

When U.S. Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

Alton J. Rahe

ALTON J. RAHE, M.S.
Project Scientist

Richard C McNee

RICHARD C. MCNEE, M.S.
Supervisor

Roy L DeHart

ROY L. DEHART
Colonel, USAF, MC
Commander

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: SAM-TR-80-46
2. GOVT ACCESSION NO.: AD-A094 899
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): ON TWO-SIDED CONFIDENCE AND TOLERANCE LIMITS FOR NORMAL DISTRIBUTIONS
5. TYPE OF REPORT & PERIOD COVERED: Final report, Jan 1967 - May 1967
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Alton J. Rahe, M.S.
8. CONTRACT OR GRANT NUMBER(s):
9. PERFORMING ORGANIZATION NAME AND ADDRESS: USAF School of Aerospace Medicine (BRA), Aerospace Medical Division (AFSC), Brooks Air Force Base, Texas 78235
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62202F, 7930-15-02
11. CONTROLLING OFFICE NAME AND ADDRESS: USAF School of Aerospace Medicine (BRA), Aerospace Medical Division (AFSC), Brooks Air Force Base, Texas 78235
12. REPORT DATE: December 1980
13. NUMBER OF PAGES: 78
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): UNCLASSIFIED

15a. DECLASSIFICATION DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report): Approved for public release; distribution unlimited.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number):
    - Confidence limits
    - Simultaneous confidence limits
    - Tolerance limits
    - Simultaneous tolerance limits
20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

This report gives known theorems on which the concept and construction of confidence and two types of tolerance limits for normal distributions are based. Procedures are presented for computing two-sided confidence and tolerance limits for means and simple linear regression data (simultaneous and nonsimultaneous limits for each type). A numerical simple linear regression example is presented showing the six types of limits. An additional bibliography is given for reference on confidence and tolerance limits when information other than what is given in the report is desired.

SECURITY CLASSIFICATION OF THIS PAGE(*When Data Entered)*

SECURITY CLASSIFICATION OF THIS PAGE(*When Data Entered)*

PREFACE

This report, with minor changes, is a thesis presented as partial fulfillment of the requirements for the Master of Science Degree in Statistics at Virginia Polytechnic Institute in 1967. Since this thesis is continually used as a source of information within the USAF School of Aerospace Medicine, it is being submitted for publication as a SAM-TR.

The author expressed appreciation to Dr. Klaus Hinkelmann and Dr. Raymond Myers, of Virginia Polytechnic Institute, for their invaluable guidance and advice to this thesis.

CONTENTS

CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

ON TWO-SIDED CONFIDENCE AND TOLERANCE LIMITS FOR NORMAL DISTRIBUTIONS

I. INTRODUCTION

In many cases of statistical inference it is more meaningful and informative to construct confidence intervals for parameters under investigation rather than to make tests of hypotheses. This requires some understanding of the concept of confidence intervals. Coupled with the understanding of confidence intervals is the understanding of tolerance limits. Frequently one finds that confidence limits are used when tolerance limits should be used, or confidence limits are computed with the general interpretation of tolerance limits.

In this report confidence limits and two types of tolerance limits are described for normal distributions giving some theorems on which the concept and construction of these limits are based. Differences and similarities between the three types of limits are pointed out. Procedures are presented for computing two-sided confidence and tolerance limits for means and for simple linear regression data (simultaneous and non-simultaneous limits for each type). For comparative purposes, the six different types of limits are computed on a numerical regression problem.

Finally, an additional bibliography is included for reference on confidence and tolerance limits when information other than what is given in the paper is desired.

II. CONFIDENCE LIMITS

Suppose a random sample of n observations ($Y_1, Y_2, \ldots, Y_n$) is drawn from a normal population in an attempt to obtain some information about the mean of the population, μ. A point estimate of the parameter μ is the sample mean, $\bar{Y}$. Although the estimate is unbiased it is not very meaningful without some measure of the possible error. Thus, frequently one determines an upper and a lower limit or a confidence interval which is rather certain to contain μ.

The general method of construction of confidence limits is as follows (4). Suppose one has a family of populations each with a known density function $p(y:\varphi)$, y being the random variable and φ the parameter in question. Suppose one has an estimator g to estimate φ, where g is a function of the observed y, and suppose that one can derive the density function of g, $p(g:\varphi)$. Now if one assumes that φ equals some particular value, say φ', then this value can be inserted and the density function $p(g:\varphi')$, the distribution of g under this assumption, can be obtained.

Under the assumption $\varphi = \varphi'$, there will be a P_1 point for the distribution of g, say g_1, which will be determined by

$$Pr\left[g \le g_1 : \varphi = \varphi'\right] = \int_{-\infty}^{g_1} p(g:\varphi')\, dg = P_1.$$

Likewise, under the same assumption there will be a P_2 point for the distribution of g, say g_2, determined by

$$\Pr\left[g \geq g_2 : \varphi = \varphi'\right] = \int_{g_2}^{\infty} p(g : \varphi')\, dg = 1 - P_2 \ . \qquad (2.1)$$

The area under the density function below g_2 is equal to P_2, and the area between g_1 and g_2 is then equal to $(P_2 - P_1) = \gamma$, say.

Now, if the value of φ' is changed, the corresponding values of g_1 and g_2 are changed. Therefore g_1 and g_2 can be regarded as functions of φ, say $g_1(\varphi)$ and $g_2(\varphi)$, respectively. In principle, one can plot these functions $g_1(\varphi)$ and $g_2(\varphi)$ against φ (see Figure 1).

Now assume that the true value of φ is actually φ_o. Then $g_1(\varphi)$ and $g_2(\varphi)$ take the values $g_1(\varphi_o)$ and $g_2(\varphi_o)$, respectively, and $\Pr\left[g \leq g_1(\varphi_o)\right] = P_1$, $\Pr\left[g \geq g_2(\varphi_o)\right] = 1 - P_2$, which imply

$$\Pr\left[g_1(\varphi_o) < g < g_2(\varphi_o)\right] = P_2 - P_1 = \gamma . \qquad (2.2)$$

Now suppose that a sample observation was taken and that a numerical value of the estimate, say g_o, was computed. Then, in Figure 1, a horizontal line can be drawn parallel to the φ axis through the point g_o on the g axis. Let this line intercept the two curves $g_2(\varphi)$ and $g_1(\varphi)$ at points A and B. Project the points A and B on to the φ axis to give $\underline{\varphi}$ and $\overline{\varphi}$. One asserts that a $(P_2 - P_1)$ confidence interval for φ is

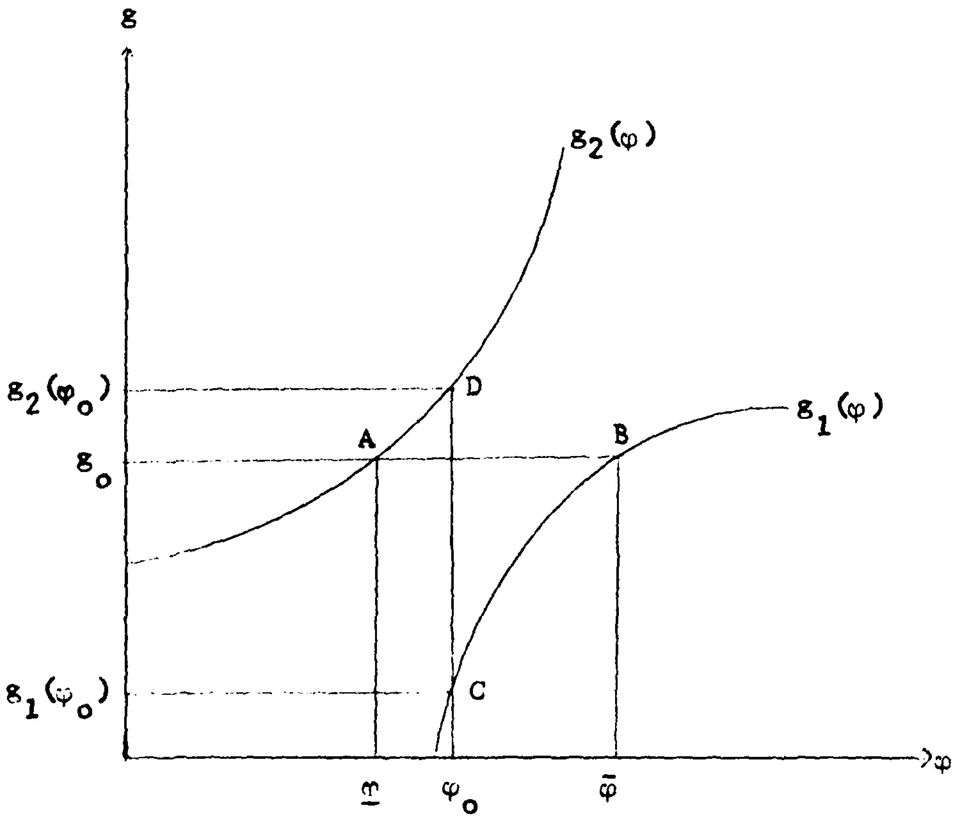

Figure 1. Plot of $g_1(\varphi)$ and $g_2(\varphi)$ Against φ for the General Method of Construction of Confidence Limits.

$(\underline{\varphi},\bar{\varphi})$, i.e.

$$\Pr\left[\underline{\varphi} < \varphi < \bar{\varphi}\right] = P_2 - P_1 = \gamma. \qquad (2.3)$$

The justification for this assertion is as follows. Enter the true value of φ_o on the φ axis; erect the perpendicular at this point to cut the curves $g_1(\varphi)$ at C and $g_2(\varphi)$ at D. At both these points φ has the values φ_o; so, at C, $g = g_1(\varphi_o)$, and, at D, $g = g_2(\varphi_o)$. The horizontal lines through C and D will intersect the g axis at $g_1(\varphi_o)$ and $g_2(\varphi_o)$, respectively. Now φ_o may be anywhere on the φ axis, but if AB intersects CD, then g_o must lie in the interval $(g_1(\varphi_o), g_2(\varphi_o))$ and simultaneously the interval $(\underline{\varphi},\bar{\varphi})$ must include φ_o. In other words, the two statements

(i) g_o lies in the interval $(g_1(\varphi_o), g_2(\varphi_o))$,

and

(ii) the interval $(\underline{\varphi},\bar{\varphi})$ includes φ_o,

are always true simultaneously or not true simultaneously. But by (2.2) the event (i) has probability (P_2-P_1); so the event (ii) must also have probability (P_2-P_1). Hence one can write

$$\Pr\left[\underline{\varphi} < \varphi_o < \bar{\varphi}\right] = P_2 - P_1 = \gamma$$

and this completes the justification of (2.3).

At the point A, the function $g_2(\varphi)$ has $\varphi = \underline{\varphi}$ and takes on the value g_o, i.e., $g_2(\underline{\varphi}) = g_o$. Now $g_2(\varphi)$ was defined as the solution of (2.1), so one can use this equation to find $\underline{\varphi}$; $\underline{\varphi}$ is obtained by solving

$$\int_{g_o}^{\infty} p(g:\varphi)\, dg = 1-P_2 = \Pr\left[g \geq g_o : \varphi = \underline{\varphi}\right]$$

Similarly, at the point B, the function $g_1(\varphi)$ has $\varphi = \overline{\varphi}$ and takes the value g_o; so $g_1(\overline{\varphi}) = g_o$ and $\overline{\varphi}$ can be found as the solution f

$$\int_{-\infty}^{g_o} p(g:\varphi)\, dg = P_1 = \Pr\left[g \leq g_o ; \varphi = \overline{\varphi}\right]$$

To determine, for instance, confidence intervals for the population mean μ, one must seek a random variable which depends on μ, no other unknown parameters, and the sample random variables, whose distribution is known. For the normally distributed variable with σ unknown the quantity

$$t = \frac{(\bar{Y}-\mu)\sqrt{n}}{s}$$

is such a random variable having Student's-t distribution with n-1 degrees of freedom (df), where

$$s = \sqrt{\frac{n \sum_{i=1}^{n} Y_i^2 - (\sum_{i=1}^{n} Y_i)^2}{n(n-1)}},$$

s^2 being an unbiased estimate of σ^2.

Before proceeding with the derivation of the confidence interval, we shall recall the definition of Student's-t distribution (5). A random variable has Student's-t distribution with n-1 df if it has the same distribution as the quotient $(u\sqrt{n-1})/v$, where u and v are independent random variables, u having a normal distribution with mean 0 and standard deviation 1, and v^2 having a chi-square (χ^2) distribution with n-1 df. More precisely, $((\bar{Y}-\mu)\sqrt{n})/\sigma$ is normally distributed with mean 0 and variance 1, and s^2/σ^2 is distributed (independently) as $\chi^2/n-1$ with n-1 df.

From tables of the Student's-t distribution one determines two percentiles, $t_{(1-\gamma)/2,n-1}$ and $t_{(1+\gamma)/2,n-1}$, such that*

$$\Pr\left[t_{(1-\gamma)/2,n-1}<t<t_{(1+\gamma)/2,n-1}\right] = \int_{t_{(1-\gamma)/2,n-1}}^{t_{(1+\gamma)/2,n-1}} f(t,n-1)\,dt = \gamma$$

where

$$f(t,n-1) = \frac{\Gamma(\frac{n}{2})(1+\frac{t^2}{n-1})^{-\frac{n}{2}}}{\sqrt{\pi(n-1)}\,\Gamma(\frac{n-1}{2})} .$$

---

*In hypothesis testing one rejects the hypothesis that $\mu = \mu_o$ if t falls outside this interval where the alternate hypothesis is that $\mu \neq \mu_o$. This represents a test of size $1-\gamma$.

Or, more precisely,

$$\Pr\left[t_{(1-\gamma)/2,n-1} < \frac{(\bar{Y}-\mu)\sqrt{n}}{s} < t_{(1+\gamma)/2,n-1}\right] = \gamma.$$

This inequality is then converted to

$$\Pr\left[\bar{Y}-t_{(1+\gamma)/2,n-1}\frac{s}{\sqrt{n}} \le \mu \le \bar{Y}-t_{(1-\gamma)/2,n-1}\frac{s}{\sqrt{n}}\right] = \gamma. \quad (2.4)$$

This interval, a confidence interval for μ, is given in most standard statistical texts (16). Owing to the fact that Student's-t distribution is symmetric, $t_{(1-\gamma)/2,n-1} = -t_{(1+\gamma)/2,n-1}$. This fact will be used throughout the remainder of the paper.

For the case where σ is known one can use (2.4) for the computation of the confidence interval by simply replacing s by σ and using for df = ∞, $t_{(1+\gamma)/2,\infty} = Z_{(1+\gamma)/2}$, the $(1+\gamma)/2$ normal deviate, since Student's-t distribution approaches the normal distribution for large degrees of freedom.

The interpretation of confidence limits is as follows. If many samples of size n were drawn from the same population and $100\gamma\%$ upper and lower limits were determined from each sample, then one would expect $100\gamma\%$ of these "random intervals" to cover the point, μ. Or, if an experimenter asserts *a priori* that an interval includes the parameter, μ, he should be making a correct statement $100\gamma\%$ of the time. In practice, one usually has only one sample from which to determine an interval estimate.

One should remember in the above discussion and throughout the rest of the paper, that upper and lower limits are computed but that frequently it is more convenient to speak of the interval formed by the limits.

Moment generating functions may be used to show that a linearly transformed normal random variable is normally distributed and that any linear combination of independent normal random variables has a normal distribution (5). The following general procedure (Procedure A) may then be used for the computation of confidence limits on any parameter or linear function of parameters φ from normal populations [e.g. $\varphi = \mu$, $\varphi = \mu_1 - \mu_2$ or $\varphi = \beta$*]:

Procedure A

1. Obtain an estimator g of φ

   e.g. $g = \bar{Y}$, $g = \bar{Y}_1 - \bar{Y}_2$, or $g = b$**

---

*population regression coefficient

** $$b = \frac{\Sigma X_i Y_i - (\Sigma X_i)(\Sigma Y_i)/n}{\Sigma X_i^2 - (\Sigma X_i)^2/n} = \frac{Sxy}{Sx^2}$$

where $\Sigma = \sum_{i=1}^{n}$

2.* Obtain the variance of g and write it in the form σ^2/n'

e.g. var $\bar{Y} = \sigma^2/n$, var$(\bar{Y}_1-\bar{Y}_2)^{**} = (\frac{1}{n_1} + \frac{1}{n_2})\,\sigma^2$,

or var(b) $= \sigma^2/Sx^2$

3. Obtain an unbiased estimate of σ^2 (usually called s^2)

e.g.

$$s^2 = \frac{\Sigma Y_i^2 - (\Sigma Y_i)^2/n}{n-1} = \frac{Sy^2}{n-1}$$

$$s^2 = \frac{Sy_1^2 + Sy_2^2}{n_1+n_2-2}$$

or

$$s^2 = \frac{Sy^2 - (Sxy)^2/Sx^2}{n-2}$$

4. Confidence interval estimate for φ^{***}=

$g \pm t_{(1+\gamma)/2,f}\sqrt{1/n'}\ s$

where $t_{(1+\gamma)/2,f}$ is the $(1+\gamma)/2$ percentage point of Student's-t distribution with f df (in the examples $f=n-1$, n_1+n_2-2, or $n-2$, respectively)

---

*The use of n' will be explained in the section on tolerance limits.

**Assuming that both populations have a common σ^2.

*** Remember $t_{(1-\gamma)/2,f} = -t_{(1+\gamma)/2,f}$.

III. TOLERANCE LIMITS

A. General Meaning of Tolerance Limits

Suppose a random sample of n observations ($Y_1, Y_2, \ldots, Y_n$) is drawn from a normal population with unknown mean, μ, and unknown variance, σ^2. Also suppose the experimenter is not interested in estimating μ as a single point, nor is he interested in finding confidence limits for μ. He is more concerned about predicting individual future values and would like to see limits where he can say with reasonable assurance that most of his future values will fall within. If he constructed these limits, which one calls tolerance limits, on his control data (normal range), then individual values falling outside these limits could be considered as being "abnormal" with a reasonable level of confidence.

Before proceeding to the details of two different types of tolerance limits, the following remarks are made to give the reader a better understanding of the general nature of the limits. For the moment, consider a normally distributed population with a known population mean, μ, and a known population variance, σ^2. One finds the two-sided tolerance limits which include 100P% of the population as $\mu - Z\sigma$ and $\mu + Z\sigma$ since

$$\int_{\mu-Z\sigma}^{\mu+Z\sigma} p(x)\, dx = P$$

where p(x) represents the density function of the normal distribution and Z is a numerical value which depends on the chosen value of P. Since the population parameters are known, the above statement can be made with 100% confidence, and one hardly has a statistical problem. For example, one is 100% confident that the tolerance limits, $\mu \pm 1.96\sigma$, contain the central 95% of the normal population.

Usually the parameters μ and σ^2 are not known, only the estimates $\bar{Y}$ and s^2. If μ and σ are replaced by $\bar{Y}$ and s one would get $\bar{Y} \pm 1.96s$ as limits in the above example. In repeated sampling from the same population these limits would vary about the population tolerance limits, $\mu \pm 1.96\sigma$, and for some samples the limits would include less than 95% of the population and for other samples more than 95%. To be reasonably sure that 100P% of the population lie between the sample tolerance limits one must find a value $k>Z$ such that there is a good chance that $\bar{Y} \pm ks$ will include 100P% of the population.

Two types of tolerance limits will be discussed: tolerance limits without confidence probability $\left[(P)TL\right]$, and tolerance limits with confidence probability $\left[(P,\gamma)TL\right]$.

B. Tolerance Limits Without Confidence Probability [(P)TL]

The problem here is to determine k so that for repeated samples of size n the <u>average</u> proportion in $\bar{Y}_i \pm ks_i$ $(i=1,2,\ldots)$ is equal to P. Wilks (20) first determined such a k, but the proof given in this paper is the proof by I.R. Savage found in an article by Proschan (14).

Let us consider as tolerance limits L_1 and L_2 the quantities $\bar{Y} \pm ks$ (two-sided limits). The proportion P' of the normal population between these limits is

$$P' = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{\bar{Y}-ks}^{\bar{Y}+ks} e^{-(Y-\mu)^2/2\sigma^2}\, dY$$

We wish to determine k so that $E(P') = P$, where

$$E(P') = \int_{-\infty}^{\infty} \int_{0}^{\infty} P' f(\bar{Y},s)\, ds\, d\bar{Y}$$

and $f(\bar{Y},s)$ is the distribution of $\bar{Y}$ and s given by

$$\frac{\sqrt{n}\,(n-1)^{(n-1)/2}\, s^{n-2}\, e^{-\left[n(\bar{Y}-\mu)^2+(n-1)s^2\right]/2\sigma^2}}{2^{\frac{n}{2}-1}\,\sigma^n \sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)}$$

Using the linear transformation, $Z = (Y-\mu)/\sigma$, $E(P')$ can be written as

$$E(P') = c_1 \int_0^{\infty} \int_{-\infty}^{\infty} \int_{\bar{Y}-ks}^{\bar{Y}+ks} e^{-Z^2/2}\, dZ\ s^{n-2}\ e^{-\left[n\bar{Y}^2+(n-1)s^2\right]/2} d\bar{Y}\ ds$$

where $c_1 = \dfrac{\sqrt{n}\ (n-1)^{(n-1)/2}}{\sqrt{2\pi}\ 2^{\frac{n}{2}-1}\ \sigma^n\ \sqrt{\pi}\ \Gamma(\frac{n-1}{2})}$ (free of k).

The conditions for differentiating under the integral hold and thus by Leibniz's rule one has

$$\frac{\partial E(P')}{\partial k} = c_1 \int_0^{\infty} \int_{-\infty}^{\infty} \left[s e^{-(\bar{Y}+ks)^2/2} + s e^{-(\bar{Y}-ks)^2/2} \right] s^{n-2} \cdot e^{-\left[n\bar{Y}^2+(n-1)s^2\right]/2}\ d\bar{Y}\ ds$$

$$= c_1 \int_0^{\infty} \int_{-\infty}^{\infty} e^{-\left[(\sqrt{n+1}\ \bar{Y}+(ks/\sqrt{n+1}))^2+(n-1+k^2n/(n+1))s^2\right]/2} \cdot s^{n-1}\ d\bar{Y}\ ds$$

$$+ c_1 \int_0^{\infty} \int_{-\infty}^{\infty} e^{\left[(\sqrt{n+1}\bar{Y}-(ks/\sqrt{n+1}))^2+(n-1+k^2n/(n+1))s^2\right]/2} \cdot s^{n-1}\ d\bar{Y}\ ds$$

Let $u = \left[\sqrt{n+1}\ \bar{Y} \pm (ks/\sqrt{n+1})\right]$, then

$$\frac{\partial E(P')}{\partial k} = c_1 \int_0^{\infty} \int_{-\infty}^{\infty} e^{-u^2/2} \frac{du}{\sqrt{n+1}} s^{n-1} e^{-\left[n-1+k^2 n/(n+1)\right] s^2/2} ds$$

$$+c_1 \int_0^{\infty} \int_{-\infty}^{\infty} e^{-u^2/2} \frac{du}{\sqrt{n+1}} s^{n-1} e^{-\left[n-1+k^2 n/(n+1)\right] s^2/2} ds$$

$$\frac{\partial E(P')}{\partial k} = c_2 \int_0^{\infty} s^{n-1} e^{-\left[n-1+k^2 n/(n+1)\right] s^2/2} ds$$

Let $y = \left[n-1+k^2 n/(n+1)\right] s^2/2$,

$$\frac{\partial E(P')}{\partial k} = c_2 \int_0^{\infty} 2^{(n-2)/2} y^{(n-2)/2} e^{-y} / \left[n-1+k^2 n/(n+1)\right]^{n/2} dy$$

$$= c_3 \frac{1}{\left[n-1+k^2\left(\frac{n}{n+1}\right)\right]^{n/2}}$$

Hence $E(P') = c_3 \int_{k_1}^{k_2} \frac{dk}{\left[n-1+k^2\left(\frac{n}{n+1}\right)\right]^{n/2}}$

where k_1 and k_2 are to be chosen so that the integral is equal to P. Let

$$t = k\sqrt{\frac{n}{n+1}}$$

so that $E(P') = c_4 \int_{t_1}^{t_2} \frac{dt}{(n-1+t^2)^{n/2}}$

$$= c_5 \int_{t_1}^{t_2} \frac{dt}{\left[1+\frac{t^2}{n-1}\right]^{n/2}}$$

But the integrand is essentially Student's-t density function with n-1 df, and when k_1 and $k_2 = -\infty$ and $+\infty$, respectively, $E(P') = 1$. Hence c_5 must be identical to the constant of Student's-t distribution. Hence for $E(P') = P$ it follows that $t_1 = t_{(1-P)/2,n-1}$ and $t_2 = t_{(1+P)/2,n-1}$. Since $t_{(1-P)/2,n-1} = -t_{(1+P)/2,n-1}$, $k = \pm t_{(1+P)/2,n-1}\sqrt{\frac{n+1}{n}}$ for tolerance limits symmetric about $\bar{Y}$.

The interval estimates

$$\bar{Y}_i \pm t_{(1+P)/2,n-1}\sqrt{\frac{n+1}{n}}\, s_i \qquad (3.1)$$

which, on the average, include 100P% of the population are referred to as tolerance limits without confidence probability or in this paper simply as (P)TL. Thus, when many samples of the same size are taken from the population and a (P)TL is calculated each time (same P), these intervals will on the average include 100P% of the population. If the experimenter asserts a priori that an interval estimate contains 100P% of the population, he stands a good chance that the interval contains in the neighborhood of 100P%, but his estimate may include considerably more or considerably less than the desired 100P%. All one does know is that the average of many of such interval estimates (expected value) contains 100P% of the population.

At this point it is not easy to see how one could generalize the above result in order to compute a (P)TL for any variate for which there is a normally distributed estimate of the mean with variance σ^2/n' and the estimate of the variance is independently distributed as $\sigma^2\chi^2/f$ with f df. The approach one can use in generalizing the procedure will be shown in the next section when considering the similarity between confidence limits and (P)TL (see page 31).

C. Tolerance Limits With Confidence Probability [(γ,P)TL]

For many situations the above tolerance interval estimate is not too useful without some measure of the possible error associated with it. Another factor which may disturb some experimenters about the (P)TL is that per interval estimate one has little assurance of always containing 100P% or more of the population. Thus, tolerance limits with confidence probability came into being. In this paper these tolerance limits will be referred to as (γ,P)TL, based on the notation in (8)*.

The problem is to find that value of k in

$$A = \frac{1}{\sigma\sqrt{2\pi}} \int_{g-ks}^{g+ks} e^{-\frac{(g-\varphi)^2}{2\sigma^2}} dg$$

such that $Pr\left[A \geq P\right] = \gamma$. A is the proportion of the population <u>actually included</u> in a <u>given</u> interval, γ is the required confidence coefficient, and P is the proportion of the population required to be included within the limits $g \pm ks$ where g is an estimate of φ, the mean of the normal population.

Wald and Wolfowitz (17) have shown how values of k may be determined to an extremely good approximation when P and γ are specified. They considered only the case in which a random sample of n is drawn from a single normal population of unknown mean and unknown variance ($f = n-1$). Wallis (18) extended their results to cover any normally distributed variable for whose mean there is a normally

*In(8), at least a proportion $\underline{\gamma}$ of the population is asserted to lie within the tolerance limits with confidence probability $\underline{\beta}$. This notation was used in (17) and may be encountered in other texts or articles.

distributed estimate with variance σ^2/n' (Wallis called it N') and for whose variance there is an estimate independently distributed as $\sigma^2\chi^2/f$ (f not necessarily equal to n-1 where n is the sample size for estimating the mean). The n' is the effective number of observations; thus, the effective number of observations for a certain statistic which when divided into the variance of an observation, gives the variance of the statistic.

Wallis summarized the Wald-Wolfowitz derivation of tolerance factors without assuming any connection between n' and f, and the following is based on his summary.

Given a statistic g having the following characteristics:

(i) It is normally distributed

(ii) Its expected value φ is the mean of a normal population with unknown variance σ^2

(iii) It has variance equal to σ^2/n', where n' is known, and an independent estimate s^2 of σ^2 is distributed as $\sigma^2\chi^2/f$ with f degrees of freedom.

The distribution of A above is clearly independent of φ and σ, since φ merely determines the point about which g will be distributed and the variance of s is proportional to σ, so without loss of generality take $\varphi = 0$ and $\sigma = 1$ in the further computation.

$\Pr\left[A>P\right]$ depends on P, k, n' and n. To emphasize the dependence on P and k for given n' and n, let $F(P,k) = Pr(A \geq P)$. Also, denote the conditional probability of A's exceeding P for a particular value of g by $F(P,k|g)$, i.e. $F(P,k|g) = Pr\left[A \geq P|g\right]$.

If $F(P,k|g)$ is known, then $F(P,k)$ may be found by forming the product

$$\left[F(P,k|g)\right]\left[\sqrt{\frac{n'}{2\pi}}\, e^{-\frac{1}{2}n'g^2}\, dg\right],$$

which represents the probability that g will lie in an interval of length dg and that A will exceed P for given g. If one integrates out g, the result is also equal to the expectation of $F(P,k|g)$ as follows:

$$F(P,k) = \sqrt{\frac{n'}{2\pi}} \int_{-\infty}^{\infty} F(P,k|g)\, e^{-\frac{1}{2}n'g^2}\, dg = E_g F(P,k|g)$$

$F(P,k)$ can be approximated by expanding $F(P,k|g)$ in a Taylor series* at $g=0$ and taking expectations.

Since $F(P,k|g)$ is an even function of g, its odd derivatives are zero, and the Taylor expansion about $g=0$ is

$$F(P,k|g) = F(P,k|0) + \frac{g^2\partial^2 F}{2!\partial g^2} + \frac{g^4\partial^4 F}{4!\partial g^4} + \cdots \qquad (3.2)$$

with all derivatives to be evaluated at $g=0$.

*Wald and Wolfowitz show the validity of the Taylor expansion.

Taking expectations, $F(P,k) = EF(P,k|g) =$

$$F(P,k|0) + \frac{1}{2n'}\frac{\partial^2 F}{\partial g^2} + \frac{1}{8n'^2}\frac{\partial^4 F}{\partial g^4} + \cdots \qquad (3.3)$$

since the second and fourth moments of g, which is normally distributed with mean 0 and variance $1/n'$, are $1/n'$ and $3/n'^2$, respectively.

On comparing the right hand sides of (3.2) and (3.3), one sees that (3.2) will become identical with (3.3), except for terms involving the second and higher even powers of $1/n'$. Thus if one sets $g = \sqrt{1/n'}$ then

$$F(P,k|\sqrt{1/n'}) \simeq F(P,k)$$

This means that in order to obtain $F(P,k)$ one has to evaluate $F(P,k|\sqrt{1/n'})$. There is a unique value of r such that

$$\frac{1}{\sqrt{2\pi}} \int_{1/\sqrt{n'}-r}^{1/\sqrt{n'}+r} e^{-Z^2/2}\, dZ = P$$

since the left side is a monotonic increasing function of r. The r corresponds with the half length ks of an interval centered at $1/\sqrt{n'}$ for which $A = P$.

The problem is to select k large enough, in the light of the sampling distribution of s, to make the probability γ that ks will be <u>at least</u> r. Thus,

$$F(P, k\sqrt{1/n'}) = Pr(s \geq r | k) = Pr(\chi_f^2 \geq fr^2 | k^2) = \gamma$$

since $\chi_f^2 = fs^2/\sigma^2$ and here $\sigma = 1$. This probability can be evaluated from tables of the chi-square distribution, after first finding r from tables of the normal distribution using a trial and error method or Newton's method (19).

After P and γ are given, one solves for k in

$\chi^2_{1-\gamma,f} = fr^2/k^2$, where $\chi^2_{1-\gamma,f}$ is that number for which $Pr\left[\chi_f^2 \geq \chi^2_{1-\gamma,f}\right] = \gamma$; then $k = ru$ where $u = \sqrt{f/\chi^2_{1-\gamma,f}}$.

The interpretation of these limits is as follows. When many random samples of the same size are taken from the normal population and a (γ,P)TL is calculated each time, then in 100γ% of the cases these limits will include <u>at least</u> 100P% of the population.

The following procedure (Procedure B) may be used to compute (γ,P)TL for any variate for which there is a normally distributed estimate of the mean with variance σ^2/n' and an estimate of the variance independently distributed as $\sigma^2\chi^2/f$ with f df:

Procedure B

1. Obtain an estimate g of the population mean (e.g. $g = \bar{Y}$, $g = \bar{Y}_1 - \bar{Y}_2$)

2. Obtain var(g) and write it in the form σ^2/n'
(e.g. $\mathrm{var}(\bar{Y})=(\frac{1}{n})\sigma^2$, $\mathrm{var}(\bar{Y}_1-\bar{Y}_2)^{*}=(\frac{1}{n_1}+\frac{1}{n_2})\sigma^2$)
3. Obtain an unbiased estimate of σ^2 (usually called s^2, with f df)
4. Decide on reasonable values of γ and P
5. Compute r:

$$r = Z_{(1+P)/2}\left[1 + \frac{1}{2n'} - \frac{2Z^2_{(1+P)/2} - 3}{24n'^2}\right],$$

from Bowker (2), where $Z_{(1+P)/2}$ is the $(1+P)/2$ percentage point of the standard normal distribution
6. Compute u:

$u^{**} = \sqrt{f/\chi^2_{1-\gamma,f}}$ where $\chi^2_{1-\gamma,f}$ is that percentile of the χ^2-distribution with f df which will be exceeded by chance $100\gamma\%$ of the time.

---

*Assuming that both populations have a common variance σ^2.

**Dixon and Massey (6) give $\sqrt{F_{1-\gamma,\infty,n-2}}$ in place of u. However the $F_{1-\gamma,\infty,n-2}$ should read $F_{\gamma,\infty,n-2}$ for the appropriate value from their table of percentiles of the $F(\nu_1,\nu_2)$ distributions. The n-2 is associated with the degrees of freedom for error in their regression procedure.

7. Compute $k = ru$

8. $(\gamma,P)TL = g \pm k\sqrt{s^2}$

Step 8 would be modified to read as $g \pm k\sqrt{s^2/m}$ if the experimenter were interested in $(\gamma,P)TL$ for future <u>means</u> based on m observations each (7).

Tabular values were obtained for r and u by Weissberg and Beatty (19), and their values are also given in Owen's <u>Handbook of Statistical Tables</u> (12). The tabulated values for r were prepared for a sample of size n from a single population and are given as $r \sim r(n,P)$. One needs to let $n = n'$ when using these tables.

Bowker (2) has shown that for large n' the expression $Z_{(1+P)/2}\left[1 + 1/2n'\right]$ may be used for r instead of the expression given in Step 5.

Bowker (3) has tabulated values of k for the special case where $f = n-1$.

Situations may arise where μ or σ is known. In the event that μ is known and σ is unknown one can use the above result as $k = Z_{(1+P)/2}\, u$ where $Z_{(1+P)/2}$ is the $(1+P)/2$ percentile point of the standard normal distribution. If σ is known and μ is unknown then the above result is used with ∞ degrees of freedom ($f = \infty$). The u will become 1, and $k = r$ which depends only on n' and P. Regardless of what level of γ is chosen u is always equal to one in the case where σ is known.

IV. RELATIONSHIP BETWEEN THE VARIOUS LIMITS

A. Contrasts of the Limits

Figure 2 gives an oversimplified comparison between the confidence limits, and the tolerance limits [(P)TL and (γ,P)TL] for different sample sizes. The "picture" was drawn as simply as possible to illustrate the basic concepts, but the following shortcomings should be realized:

1. At each sample size (except $n=\infty$), each interval is an estimate and is not necessarily symmetric about μ.
2. At each sample size (except $n=\infty$), one should visualize many confidence interval estimates with 100γ% of them covering μ, many (P)TL estimates whose average interval covers 100P% of the population, and many (γ,P)TL with 100γ% of these intervals covering <u>at least</u> 100P%.
3. When σ is not known, all estimates mentioned in 2 (above) will usually be of unequal length.

The (P)TL gives an estimate of the <u>interval</u> $\mu \pm k\sigma$ in the same manner as $\bar{Y}$ gives an estimate of the <u>point</u> μ. The (γ,P)TL are in nature comparable to the confidence limits because these tolerance limits give a "confidence interval" about an <u>interval</u> (including at least 100P% of the population), while the confidence limits give a confidence interval about a <u>point</u>.

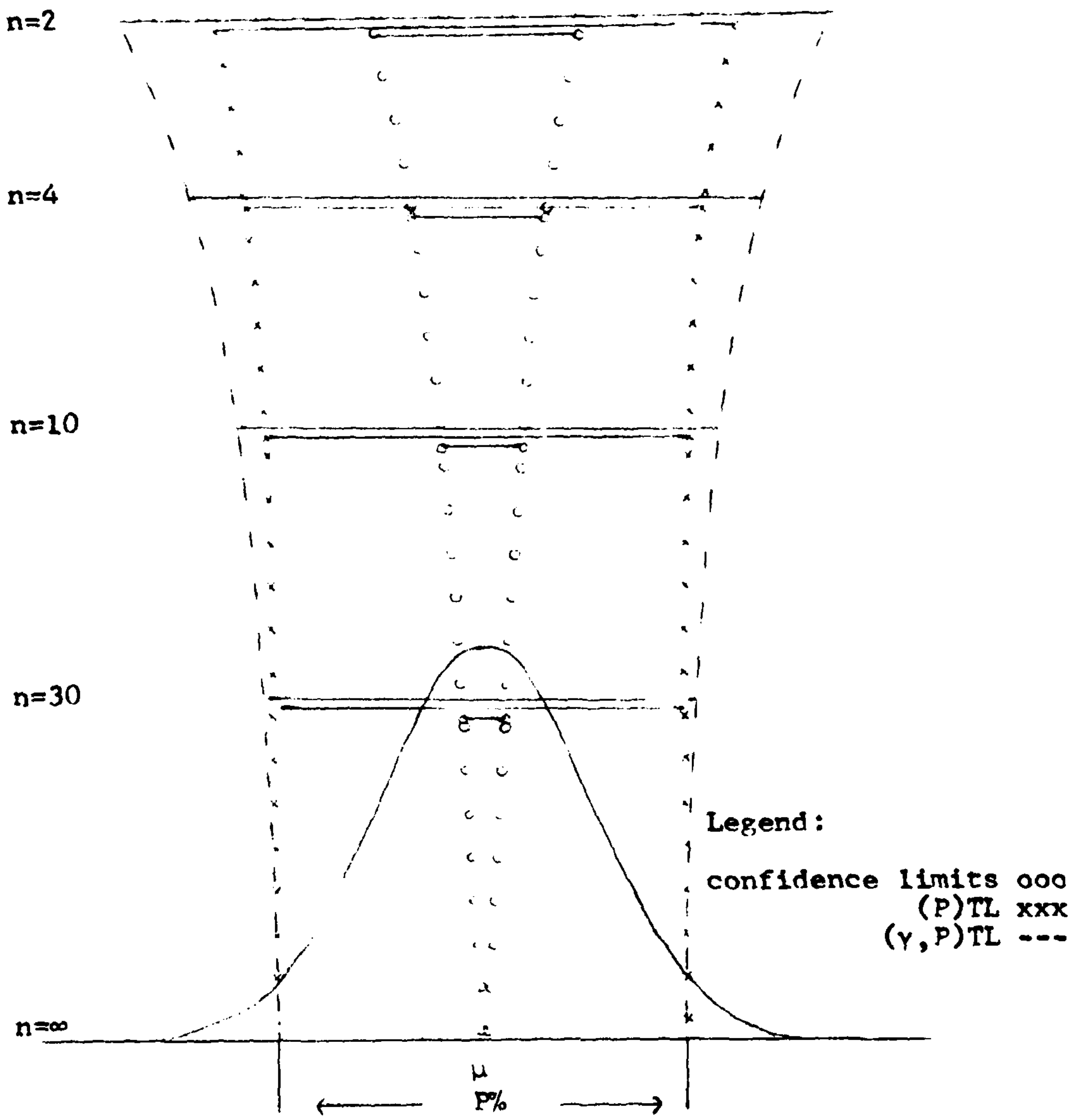

Figure 2. Oversimplified Comparison Between Confidence Limits, (P)TL, and (γ,P)TL on a Simple Mean for Different Sample Sizes.

For a very large sample the confidence limits converge to one point, the parameter (see Figure 2). This can easily be verified from the previous formulas. As sample size and degrees of freedom increase for the normal distribution the (γ,P)TL and the (P)TL approach essentially two limiting parameters with 100% confidence including the proportion P of the population.

B. Similarity Between Confidence Limits and Tolerance Limits [(P)TL]

The following is based on Proschan's article. Frequently, experimenters are interested in finding a prediction (or "confidence") interval for an additional observation from the same population. Most standard statistical texts (16) show that

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\left[\frac{\left[\Sigma^* Y_1^2 - (\Sigma Y_1)^2/n_1\right] + \left[\Sigma Y_2^2 - (\Sigma Y_2)^2/n_2\right]}{n_1+n_2-2}\right]\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

is distributed as Student's-t with $f = n_1+n_2-2$. One may now use this relationship to find the following prediction interval for the value of one additional observation $Y_2 (n_2=1)$:

---

* All $\Sigma = \sum_{i=1}^{n_1}$ or $\sum_{i=1}^{n_2}$

$$\Pr\left[\bar{Y}_1 - t_{(1+\gamma)/2,n_1-1}\sqrt{(n_1+1)/n_1}\; s_1 < Y_2 < \bar{Y}_1 + t_{(1+\gamma)/2,n_1-1}\sqrt{(n_1+1)/n_1}\; s_1\right] = \gamma \quad (4.1)$$

where

$$s_1 = \sqrt{\frac{\Sigma Y_1^2 - (\Sigma Y_1)^2/n_1}{n_1-1}} \quad .$$

This simply means that if pairs of samples of size n_1 and 1 for $\bar{Y}_1$ and Y_2, respectively, are drawn repeatedly, then $100\gamma\%$ of the Y_2's will lie in the above interval. It does not mean that if one sample of size $n_1(\bar{Y}_1)$ were drawn, to be followed by the drawing of many additional Y_2's that $100\gamma\%$ of these Y_2's will lie in the interval.

Notice that the $100\gamma\%$ confidence limits for the value of one additional observation (4.1) is the same as the (P)TL (3.1) except for the subscript on t, remembering that $t_{(1-P)/2,n-1} = - t_{(1+P)/2,n-1}$. How is this confidence or prediction interval related to the (P)TL ? An intuitive explanation of their relationship may go as follows. The $\bar{Y}_1 \pm t_{(1+\gamma)/2,n_1-1}\sqrt{(1/n)+1}\; s_1$ in (4.1) is an estimate of $\mu \pm t_{(1+\gamma)/2,\infty}\sqrt{1}\,\sigma$, and substituting, (4.1) would become

$$\Pr\left[\mu - t_{(1+\gamma)/2,\infty}\,\sigma < Y_2 < \mu + t_{(1+\gamma)/2,\infty}\,\sigma\right] = \gamma .$$

This interval is fixed and contains the central $100\gamma\%$ of the future Y_2's from the population. Thus each (4.1) is an estimate of an interval which contains $100\gamma\%$ of the population. However, this is the definition of (P)TL in Section III, replacing γ with P. Hence, confidence limits with confidence coefficient γ for a second sample of size one are identical with tolerance limits that will include a proportion P on the average.

Paulson (13) proves the following simple lemma on the relationship between confidence limits (γ) for a future random observation and (P) tolerance limits: If confidence limits $U_1(x_1,\ldots,x_n)$ and $U_2(x_1,\ldots,x_n)$ on a probability level $= \gamma$ are determined for g, a function of a future sample of k observations, and

$$P = \int_{u_1}^{u_2} \Psi(g)\, dg,$$

then $E(P) = \gamma$. Let $\Psi(g)\, dg$ and $\phi(U_1,U_2)\, dU_1\, dU_2$ denote the distribution of g and U_1, U_2 respectively, then by the definition of expected value

$$E(P) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\int_{u_1}^{u_2} \Psi(g)\, dg \right] \phi(U_1,U_2)\, dU_1\, dU_2 \; .$$

This triple integral is, however, exactly the probability that g will lie between U_1 and U_2, which by the nature of confidence limits must equal γ, which proves the lemma.

Following the procedure of computing confidence limits for the next observation, one can quite easily compute (P)TL for any variate for which there is a normally distributed estimate of the mean with variance σ^2/n' and the estimate of the variance is independently distributed as $\sigma^2\chi^2/f$ with f df. For example, the (P)TL for Y_1-Y_2 when given n_1 observations from the Y_1 population and n_2 observations from the Y_2 population is obtained from

$$\Pr\left[t_{(1-P)/2} \le \frac{(\bar{Y}_1-\bar{Y}_2) - (Y_1-Y_2)}{\sqrt{s^2(\frac{1}{n_1}+\frac{1}{n_2}+1+1)}} \le t_{(1+P)/2}\right] = P$$

where s^2 is the pooled sample variance. This expression is then rearranged as follows:

$$\Pr\left[(\bar{Y}_1-\bar{Y}_2)+t_{(1-P)/2}\sqrt{s^2(\frac{1}{n_1}+\frac{1}{n_2}+2)} \le Y_1-Y_2 \le (\bar{Y}_1-\bar{Y}_2) + t_{(1+P)/2}\sqrt{s^2(\frac{1}{n_1}+\frac{1}{n_2}+2)}\right] = P$$

A summary of the computing procedures for the two-sided confidence limits and both types of tolerance limits on normal populations is given in Table 1.

TABLE 1. COMPUTATIONAL PROCEDURES OF CONFIDENCE LIMITS, (P)TL, AND (γ,P)TL FOR NORMAL POPULATIONS

Source	Parameters	Step # 1	Step # 2
Confidence Limits	φ unknown(U) σ^2U	Obtain estimate g of φ	Obtain var(g) = σ^2/n'
"	φ U σ^2known(K)	"	"
(P)TL	φ U σ^2U	"	Var. g + var. of future <u>single</u> (g)
"	φ U σ^2K	"	"
"	φ K σ^2U	-	"
"	φ K σ^2K	-	"
(γ,P)TL	φ U σ^2U	Obtain estimate g of φ	Obtain var(g) = σ^2/n'
"	φ U σ^2K	"	"
"	φ K σ^2U	-	"
"	φ K σ^2K	-	"

(Table 1 continued.)

Source	Parameters	Step # 3	Step # 4
Confidence Limits	φ unknown(U) σ^2U	Obtain estimate of σ^2(called s^2)	Decide on γ
"	φ U σ^2known(K)	-	"
(P)TL	φ U σ^2U	Obtain estimate of σ^2(called s^2)	Decide on P
"	φ U σ^2K	-	"
"	φ K σ^2U	Obtain estimate of σ^2(called s^2)	"
"	φ K σ^2K	-	"
(γ,P)TL	φ U σ^2U	Obtain estimate of σ^2(called s^2)	Decide on γ and P
"	φ U σ^2K	-	Decide on P only
"	φ K σ^2U	Obtain estimate of σ^2(called s^2)	Decide on γ and P
"	φ K σ^2K	-	Decide on P only

(Table 1 continued.)

Source	Parameters	Step #5*
Confidence Limits	φ unknown(U) σ^2U	Confidence interval of $\varphi = g \pm t_{(1+\gamma)/2,f}\sqrt{s^2/n'}$
"	φ U σ^2known(K)	$g \pm t_{(1+\gamma)/2,\infty}\sqrt{\sigma^2/n'}$
(P)TL	φ U σ^2U	(P)TL = $g \pm t_{(1+P)/2,f}\sqrt{s^2(1/n' + 1)}$ **
"	φ U σ^2K	$g \pm t_{(1+P)/2,\infty}\sqrt{\sigma^2(1/n' + 1)}$
"	φ K σ^2U	$\varphi \pm t_{(1+P)/2,f}\sqrt{s^2(1/n' + 1)}$
"	φ K σ^2K	$\varphi \pm t_{(1+P)/2,\infty}\sqrt{\sigma^2(1/n' + 1)}$
(γ,P)TL	φ U σ^2U	$r = t_{(1+P)/2,\infty}\left[1 + \frac{1}{2n'} \quad \frac{2t^2_{(1+P)/2,\infty} - 3}{24(n')^2}\right]$
"	φ U σ^2K	"
"	φ K σ^2U	$r = t_{(1+P)/2,\infty}$
"	φ K σ^2K	"

* and ** see page 39

(Table 1 continued.)

Source	Parameters	Step #6	Step #7	Step #8
Confidence Limits	φ unknown (U) σ^2 U			
"	φ U σ^2 known (K)			
(P)TL	φ U σ^2 U			
"	φ U σ^2 K			
"	φ K σ^2 U			
"	φ K σ^2 K			
(γ,P)TL	φ U σ^2 U	$u=\sqrt{\frac{f}{\chi^2_{1-\gamma,f}}}$***	$k=ru$	$(\gamma,P)TL = g \pm k\sqrt{s^2}$
"	φ U σ^2 K	1	$k=r$	$g \pm k\sqrt{\sigma^2}$
"	φ K σ^2 U	$u=\sqrt{\frac{f}{\chi^2_{1-\gamma,f}}}$	$k=ru$	$\varphi \pm k\sqrt{s^2}$
"	φ K σ^2 K	1	$k=r$	$\varphi \pm k\sqrt{\sigma^2}$

***See page 39

(Table 1 continued.)

* $t_{\lambda,f}$ is the λ percentage point of Student's-t distribution with f df.

** Formula as given is not always correct depending on the φ under consideration. See page 34.

*** $\chi^2_{1-\gamma,f}$ is the percentage point of the χ^2 distribution with f df which will be exceeded by chance $100\gamma\%$ of the time.

V. LIMITS IN SIMPLE LINEAR REGRESSION

A. Background

In linear regression, Y values are obtained from several populations, each population being determined by a corresponding X value. The X variable is fixed or measured without error. The following assumptions are usually made about the "true" model:

1. The distribution of Y for each X is normal.
2. The mean values of Y lie exactly on the line $\mu_{Y \cdot X} = \alpha + \beta X$.
3. The variance of Y, σ^2, is the same for each X.
4. The Y observations are statistically independent.

The classical "least squares" procedure is used for "fitting" a line which best describes the linear relationship between the (X_i, Y_i) pairs of observations. This procedure determines values of a and b which minimize

$$SSD = \sum_{i=1}^{n} (Y_i - a - bX_i)^2.$$

The b for the "fitted" line is called the regression coefficient, and the a is called the intercept. The line is called a regression line, and its equation is called a regression equation.

B. Confidence Limits

1. Non-simultaneous confidence limits

Frequently textbooks give $100\gamma\%$ confidence limits on the population mean of Y at a particular X_o value, $\mu_{Y \cdot X_o}$. The concept of computing confidence limits on a single normal population is simply applied repeatedly to the Y data at the different values of X. The intervals are not independent of each other because they all depend on the same regression line. These intervals will be referred to as non-simultaneous confidence limits (intervals).

The interpretation for any one of these populations is that if many samples of the same size were drawn from the same population of Y's at X_o and an interval were constructed for each sample, then one would expect $100\gamma\%$ of these "random intervals" to cover the fixed point $\mu_{Y \cdot X_o}$.

Procedure A for the computation of confidence limits may be used repeatedly to compute $100\gamma\%$ non-simultaneous confidence limits for different values of X (call the X under consideration, X_o). The procedure is given below for simple linear regression problems and will be referred to as Procedure C.

Procedure C

1. $\hat{Y} = a + bX_o$, where

$$b = \frac{\Sigma^{*}XY - \frac{\Sigma X\,\Sigma Y}{n}}{\Sigma X^2 - \frac{(\Sigma X)^2}{n}} = \frac{Sxy}{Sx^2}$$

and

$$a = \bar{Y} - b\bar{X}$$

2. $$\mathrm{Var}(\hat{Y}) = \sigma^2_{Y\cdot X}\left[\frac{1}{n} + \frac{(X_o-\bar{X})^2}{Sx^2}\right]$$

3. $$s_{Y\cdot X} = \sqrt{\frac{Sy^2 - (Sxy)^2/Sx^2}{n-2}}$$

where $Sy^2 = \Sigma Y^2 - (\Sigma Y)^2/n$

4. $$\mathrm{Conf.}(\mu_{Y\cdot X_o}) = \hat{Y} \pm t_{(1+\gamma)/2,f}\left[\frac{1}{n} + \frac{(X_o-\bar{X})^2}{Sx^2}\right]^{\frac{1}{2}} s_{Y\cdot X}$$

with $f = n-2$.

If each confidence limit is considered a function of X, then the limits define the two branches of a hyperbola with the fitted line as the diameter. The interval has minimum length for $X = \bar{X}$, and its length increases as $|(X-\bar{X})|$ increases.

---

* all $\Sigma = \sum_{i=1}^{n}$

2. Simultaneous confidence limits

As mentioned before, repeated use of the non-simultaneous confidence limits would result in error because of the lack of independence of the intervals. In 1929, Working and Hotelling (22) worked out a procedure whereby they found a confidence region for an entire regression line. They computed a confidence region, not an interval, which covered the whole line, not only one point on the line. This procedure later turned out to be a special case of Scheffé's simultaneous confidence intervals (15). Wilks (21) gives a proof of Scheffé's method for simultaneous confidence intervals in his text, and it is his proof that is given in this paper.

The basic result due to Scheffé is as follows:

Suppose $\underline{u}' = (u_1, \ldots, u_k)$ is a k-dimensional random variable having normal distribution

$$N(\underline{\mu}, A\sigma^2)$$

where $\underline{\mu}' = (\mu_1, \mu_2, \ldots, \mu_k)$ is the vector of the means and A is the variance-covariance matrix (non-singular) with elements a_{ij}, and σ^2 is unknown. Let S = residual sum of squares, then S/σ^2 is a random variable independent of $(u_1, \ldots, u_k)$ which follows the chi-square distribution with f df. Let $F_{\gamma,k,f}$ be the $100\gamma\%$ point of the F-distribution and let $\delta = \sqrt{(S/f)(kF_{\gamma,k,f})}$. We can then state the following theorem: If θ is the set of all real vectors $(c_1,\ldots,c_k)$ where $c_1,\ldots,c_k$ are not all zero, the inequalities

$$\sum_i c_i u_i - \delta\sqrt{\sum_{i,j} a_{ij}c_i c_j} \le \sum_i c_i \mu_i \le \sum_i c_i u_i + \delta\sqrt{\sum_{i,j} a_{ij}c_i c_j} \qquad (5.1)$$

hold simultaneously with probability γ for all $(c_1,\ldots,c_k)$ in θ.

To prove the theorem one should first note that $(\underline{u}-\underline{\mu})'A^{-1}(\underline{u}-\underline{\mu})/\sigma^2 = (1/\sigma^2)\sum_{ij} a^{ij}(u_i-\mu_i)(u_j-\mu_j)$ and S/σ^2 are independent random variables having chi-square distribution with k and f df, respectively, with a^{ij} being the elements of A^{-1}. Hence $(f/kS)\sum_{i,j} a^{ij}(u_i-\mu_i)(u_j-\mu_j)$ has F-distribution. Therefore

$$\Pr\left[\sum_{i,j} a^{ij}(u_i-\mu_i)(u_j-\mu_j) < \delta^2\right] = \gamma \qquad (5.2)$$

where $\delta^2 = (kS/f)\,F_{\gamma,k,f}$.

Next Wilks makes use of k-dimensional geometric concepts and terminology. The set of points in the space of $(\mu_1,\ldots,\mu_k)$ for which

$$\sum_{i,j} a^{ij}(u_i-\mu_i)(u_j-\mu_j) < \delta^2$$

is the interior of a $100\gamma\%$ confidence ellipsoid for the true parameter point $(\mu_1,\ldots,\mu_k)$ centered at $(u_1,\ldots,u_k)$. If one considers the set of points in the space of $(\mu_1,\ldots,$

μ_k) contained between all possible pairs of parallel (k-1)-dimensional hyperplanes tangent to this ellipsoid then this set of points constitutes the interior of the ellipsoid (5.2) and the probability associated with this set is γ.

Wilks then goes on to show that for any particular choice of $(c_1,\ldots,c_k)$ in θ the two parallel (k-1)-dimensional hyperplanes in the space of $(\mu_1,\ldots,\mu_k)$ having equations

$$\sum_i c_i\mu_i = \sum_i c_i u_i \pm \delta\sqrt{\sum_{i,j} a_{ij}c_i c_j} \qquad (5.3)$$

are tangent to the ellipsoid

$$\sum_{i,j} a^{ij}(\mu_i - u_i)(\mu_j - u_j) = \delta^2 \qquad (5.4)$$

Any point $(\mu_1,\ldots,\mu_k)$ between the two hyperplanes (5.3) satisfies (5.1). For the moment let $\mu_i - u_i = y_i$. Then (5.4) can be written as

$$\sum_{i,j} a^{ij}y_i y_j = \delta^2, \qquad (5.5)$$

and the equation of an arbitrary hyperplane in the space of $(y_1,\ldots,y_k)$ can be written as

$$\sum_i c_i y_i = d. \qquad (5.6)$$

Now one must find the two values of d for which the hyperplane (5.6) is tangent to the ellipsoid (5.5). Using a Lagrange multiplier λ, one must find the stationary points in the $(y_1,\ldots,y_k)$-space of

$$\phi = \tfrac{1}{2}\lambda(\delta^2 - \sum_{i,j} a^{ij}y_iy_j) + \sum_i c_iy_i \; .$$

Differentiating with respect to y_j one finds

$$-\lambda \sum_i a^{ij}y_i + c_j = 0 \quad \text{or}$$

$$y_i = (1/\lambda) \sum_j a_{ij} c_j \tag{5.7}$$

Substituting in (5.4) one finds

$$\lambda = \pm(1/\delta)\sqrt{\sum_{i,j} a_{ij}c_ic_j} \tag{5.8}$$

From (5.8), (5.7), and (5.6) one finds

$$d = \pm\, \delta\sqrt{\sum_{i,j} a_{ij}c_ic_j}$$

Substituting this value of d in (5.6) and using the fact that $y_i = \mu_i - u_i$, one obtains (5.3) as the equations of the two parallel tangent hyperplanes for specified $(c_1,\ldots,c_k)$. This implies (5.1) and hence proves the theorem.

In this paper one uses Scheffé's method (S-Method) of multiple comparison as stated in the preceding theorem to the family $[\alpha+\beta(X-\bar{X})]$, corresponding to the two-dimensional space $[c_1\alpha+c_2\beta]$, i.e. $c_1 = 1$ and $c_2 = X-\bar{X}$. With this procedure one can compute confidence limits for any number of different X values and say that <u>all</u> of the intervals simultaneously cover the corresponding $\mu_{Y\cdot X}$ values for $100\gamma\%$ of such random confidence regions.

The results from the S-Method show that the same procedure, Procedure C on page 41, may be used to compute these simultaneous confidence limits as was used to compute the non-simultaneous confidence limits with the following modification: In step 4, the quantity $\sqrt{2F_{\gamma,2,n-2}}$ is used instead of $t_{(1+\gamma)/2,n-2}$.

These simultaneous confidence limits also define the two branches of a hyperbola with the fitted line as the diameter. As might be expected, for a given γ level, the branches of the hyperbola for the simultaneous limits are farther apart than those for the non-simultaneous limits.

C. Non-Simultaneous Tolerance Limits

1. Non-simultaneous (P)TL

Frequently, prediction intervals are also computed for simple linear regression problems (11). The practical use of the non-simultaneous (P)TL is rather restricted since limits, like the non-simultaneous confidence limits, are not independent of each other. The same is true here as was for the confidence limits in that the concept of computing a (P)TL on a single normal population is applied repeatedly to the Y data at different values of X.

The procedure for computing non-simultaneous (P)TL is the same as Procedure C on page 41 for computing non-simultaneous confidence limits with the following <u>modification</u>: In step 2 of the procedure the variance of $\hat{Y}$ is

$$\sigma^2_{Y \cdot X}\left[1 + \frac{1}{n} + \frac{(X_o-\bar{X})^2}{Sx^2}\right]$$

which takes into consideration the variance associated with the additional observation.

These non-simultaneous (P)TL also define the two branches of a hyperbola with the fitted line as the diameter. With these limits one can rightfully say only that for <u>one</u> future X_o value 100P% of the Y values will on the average lie within the given limits.

2. Non-simultaneous (γ,P)TL

As mentioned before, the (P)TL is simply an estimate of the interval and it does not give the experimenter any assurance of including at least a desired proportion of the population. The more desirable statement would include <u>at least</u> 100P% of the population with a predetermined level of confidence (γ). Whenever textbooks consider tolerance limits in simple regression, the non-simultaneous (γ,P)TL are most frequently mentioned (1), (6).

Procedure B on page 26 is used repeatedly for different X values to compute the non-simultaneous (γ,P)TL. Again, the loci of the tolerance limits may be plotted as a hyperbola with the fitted line as diameter. It must be re-emphasized that these limits are not independent of each other and hence do not hold for different values of X simultaneously. Generally, these limits are farther apart than the non-simultaneous (P)TL when using a reasonable 100γ% confidence level.

D. Simultaneous Tolerance Limits

1. Background

Lieberman (9) first considered the joint prediction interval for the response at each of K separate values of the independent variable when all K predictions must be based upon the original fitted model. He describes three methods, one exact and two approximate. For the exact method the probability is 100γ% that all K future observations fall within their respective intervals, for the approximate methods the probability is greater than 100γ%.

These prediction regions apply only to a specified number K of future responses at each of K separate X values. However, when K is unknown and possibly arbitrarily large these results are no longer valid. A solution to the problem of arbitrary K is given in terms of simultaneous tolerance limits (intervals) on the distribution of future observations. In this paper two types of simultaneous tolerance intervals will be considered-simultaneous (P)TL and simultaneous (γ,P) TL.

2. Simultaneous (P)TL

In an attempt to overcome the limitation of the non-simultaneous (P)TL on Y at a particular X_o, simultaneous (P)TL should perhaps be considered in simple linear regression. With these simultaneous (P)TL, one may say that on the average 100P% of the Y population values are included in each interval and that this statement may be made for any number of different X values simultaneously.

The computing procedure for these simultaneous (P)TL is analogous to the computation of simultaneous confidence limits. Thus Procedure C on page 41, procedure for computation of non-simultaneous confidence limits, may be used to compute the simultaneous (P)TL with the following two modifications: In Step 2,

$$\mathrm{var}(\hat{Y}) = \sigma^2_{Y \cdot X}\left[1 + \frac{1}{n} + \frac{(X_o - \bar{X})^2}{Sx^2}\right]$$

and in Step 4, $\sqrt{2F_{\gamma,2,n-2}}$ is used instead of $t_{(1+\gamma)/2,n-2}$.

As expected, for a given P and γ, the branches of the hyperbola for the simultaneous (P)TL are farther apart than those for the non-simultaneous (P)TL.

3. Simultaneous (γ, P)TL

Each of the previously mentioned tolerance limits procedures in simple linear regression had its limitation. However, one can see that the limits for each procedure were getting wider (unfortunately), but closer to what seems, in most cases, to be in what the experimenter is actually interested. At least, each successive procedure was better than simply using non-simultaneous confidence limits and pretending that one had limits which included a given percentage of the population at some chosen level of confidence. Simultaneous (γ, P)TL appear to be the proper limits for most experimenters to use.

The approach used in the paper for the derivation of the simultaneous (γ, P)TL in regression is the simplest of four approaches presented by Lieberman and Miller (10). The authors made use of the Bonferroni inequality $P[AB] \geq 1 - P[A^c] - P[B^c]$, where A^c and B^c denote the complement of A and B, respectively. In this approach they employed the inequality to combine simultaneous confidence intervals on the regression means, as obtained by Scheffé, and the confidence interval for the standard deviation to construct a two-sided simultaneous (γ, P)TL. The two-sided confidence region for the regression line is obtained from

$$\Pr\left[|\alpha+\beta(X-\bar{X})-a-b(X-\bar{X})| \le s_{Y\cdot X}(2F_{(1+\gamma)/2,2,n-2})^{\frac{1}{2}}\sqrt{\frac{1}{n}+\frac{(X-\bar{X})^2}{Sx^2}},\right.$$

$$\left.\text{for all } X\right] = (1+\gamma)/2. \quad (5.9)$$

An upper bound on σ is obtained from a one-sided chi-square confidence interval:

$$\Pr\left[\sigma \le s_{Y\cdot X}\left[\frac{n-2}{\chi^2_{(1-\gamma)/2,n-2}}\right]^{\frac{1}{2}}\right] = \frac{1+\gamma}{2} \qquad (5.10)$$

where $\chi^2_{(1-\gamma)/2,n-2}$ is the $(1-\gamma)/2$ percentage point of the chi-square distribution for n-2 df. With use of the Bonferroni inequality the confidence statements (5.9) and (5.10) are combined into a joint confidence statement with probability greater than or equal to γ as:

$$\Pr\left[|\alpha+\beta(X-\bar{X}) \pm Z_{(1+P)/2}\,\sigma-a-b(X-\bar{X})| \le s_{Y\cdot X}\left[(2F_{(1+\gamma)/2,2,n-2})^{\frac{1}{2}}\right.\right.$$

$$\left.\left.\sqrt{\frac{1}{n}+\frac{(X-\bar{X})^2}{Sx^2}} + Z_{(1+P)/2}\left[\frac{n-2}{\chi^2_{(1-\gamma)/2,n-2}}\right]^{\frac{1}{2}}\right) \text{ for all } X, P\right] \ge \gamma$$

where $Z_{(1+P)/2}$ is the $(1+P)/2$ percentage point of the standard normal distribution.

Lieberman and Miller describe the simultaneous (γ, P) TL in simple regression, as follows: "If for a single regression line $\left[\hat{Y}=a+b(X_0-\bar{X})\right]$ one asserts that the proportion of future observations falling within the given tolerance limits (for any X), is at least P, and similar statements are made repeatedly for different regression lines $\hat{Y} = \left[a + b(X_i - \bar{X})\right]$, then for 100γ% of the different regression lines the statements will be correct". One may reword Lieberman and Miller's quotation as follows in order to give an analogous statement for the (γ,P)TL in Section III: "If for a single mean, $\bar{Y}$, one asserts that the proportion of future observations falling within the given tolerance limits is at least P, and similar statements are repeatedly for different estimates of the mean, then for 100γ% of the different estimates the statements will be correct."

The authors did not appear to have any strong preference for any one of their four procedures. They then go on to say, "The widths of these simultaneous limits (talking about the four procedures in general) vary from slightly larger to about twice as large as the non-simultaneous intervals. This gives a rough indication of the price the experimenter will have to pay, or should be paying, for simultaneity". Many experimenters may feel that these limits will be too large to be of any practical benefit. In these situations, depending on the nature of the data, the experimenter should settle for smaller P and/or smaller γ levels. Smaller or more desirable limits are not necessarily justified when obtained by a procedure which should not have been used or a procedure which gives less precise information.

The computation of the simultaneous (γ, P)TL of the form $\hat{Y} \pm k' s_{Y \cdot X}$ in simple linear regression is given in Procedure D (fixed central proportion P for all X's):

Procedure D

1. $\hat{Y} = \bar{Y} + b(X_o - \bar{X})$
2. $var(\hat{Y}) = \sigma^2_{Y \cdot X}(d)$

   where $d = \frac{1}{n} + \frac{(X_o - \bar{X})^2}{Sx^2}$
3. $s_{Y \cdot X} = \sqrt{\frac{Sy^2 - (Sxy)^2/Sx^2}{n-2}}$
4. Decide on reasonable levels of P and γ
5. $k' = \sqrt{2F_{(1+\gamma)/2,2,n-2}}\sqrt{d} + Z_{(1+P)/2}\sqrt{(n-2)/\chi^2_{(1-\gamma)/2,n-2}}$
6. $\hat{Y} \pm k' s_{Y \cdot X}$
7. Steps (1),(2),(5), and (6) should be repeated for several X values (covering the range of X's). The loci of the limits may be plotted as a hyperbola with the fitted line as diameter.

E. Regression Through the Origin

In some situations the relationship between Y and X is such that when X=0 also Y=0. Thus, one is interested in passing the regression line through the origin, and the required equation is of the type, $\mu_{Y \cdot X} = \beta X$. As in the previous case, it is assumed that deviations from the regression line are normally distributed with a common variance. Of course, the parameter estimates for this model are not the same as for the previous model, $\mu_{Y \cdot X} = \alpha + \beta X$.

The same procedure (Procedure C) for the computation of non-simultaneous confidence limits may be applied to this model as was used for the previous model using the different estimates:

1. $\hat{Y} = bX$ where $b = \dfrac{\Sigma^* X_i Y_i}{\Sigma X_i^2}$

2. $\mathrm{Var}(\hat{Y}_o) = \sigma^2_{Y \cdot X}\left[\dfrac{X_o^2}{\Sigma X_i^2}\right]$

3. $s'_{Y \cdot X} = \sqrt{\Sigma Y_i^2 - ((\Sigma X_i Y_i)^2 / \Sigma X_i^2)}$

   with n-1 degrees of freedom (f)

4. Confidence limits for $\mu_{Y \cdot X_o} = \hat{Y} \pm t_{(1+\gamma)/2, f}\left[\dfrac{X_o^2}{\Sigma X_i^2}\right]^{\frac{1}{2}} s'_{Y \cdot X}$

For $X_o = 0$ (the origin), the above procedure shows a confidence interval of 0. Initially one may feel that this is incorrect. However, for this point there is no sampling

---

* All $\Sigma = \sum_{i=1}^{n}$.

variation, the regression equation was "forced" through this point. It is easy to see that these confidence intervals increase as X_o increases. This "fan" appearance of the confidence limits is unlike the hyperbolic confidence limits obtained for the previous model.

The remainder of the confidence and tolerance intervals can be computed for $\mu_{Y \cdot X} = \beta X$ using the basic quantities given in the procedure on the previous page.

VI. NUMERICAL EXAMPLE

A summary of the computing formulas for the various confidence and tolerance limits in simple linear regression are given in Table 2. The values from the various distributions have all been given in terms of the F-distribution in this table.

A numerical example has been presented so that the reader can appreciate to a fuller extent the various computational procedures, and can graphically see the difference (if any) in the interval widths for the various procedures.

The example used in this paper is the same as the numerical example presented in Lieberman & Miller's paper using 15 hypothetical pairs of values on speed of a missile (Y) and orifice opening (X). The underlying relationship between these two variables is of the form

Expected speed (miles/hr) = α + β orifice opening (inches).

The necessary quantities from the data for the desired computations were [as given in (10)]:

$$\bar{X} = 1.3531$$

$$\bar{Y} = 5219.3$$

$$Sx^2 = \Sigma(X-\bar{X})^2 = .011966$$

$$\hat{Y} = -19{,}041.9 + 17930X$$

$$s = 130.5 \text{ with } f = 13$$

$$n = 15$$

TABLE 2. COMPUTATIONAL PROCEDURES FOR VARIOUS TYPES OF CONFIDENCE AND TOLERANCE LIMITS IN SIMPLE LINEAR REGRESSION

Source \ Step #	1	2	3
Non-simultaneous confidence limits (Procedure C)	$\hat{Y}=a+bX_o$	$\sigma^2_{Y\cdot X}\left[\frac{1}{n}+\frac{(X_o-\bar{X})^2}{Sx^2}\right]$ $=\sigma^2_{Y\cdot X}(d)$	$s_{Y\cdot X}=\sqrt{\frac{Sy^2-\frac{(Sxy)^2}{Sx^2}}{n-2}}$
Simultaneous confidence limits	"	"	"
Non-simultaneous (P)TL	"	$\sigma^2_{Y\cdot X}(1+d)$	"
Simultaneous (P)TL	"	"	"
Non-simultaneous (γ,P)TL (Procedure B)	"	$\sigma^2_{Y\cdot X}(d)$	"
Simultaneous (γ,P)TL (Procedure D)	"	"	"

Notes: $a = \bar{Y}-b\bar{X}$

$$b = \frac{\Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{n}}{\Sigma X^2 - \frac{(\Sigma X)^2}{n}} = \frac{Sxy}{Sx^2}$$

$$Sy^2 = \Sigma Y^2 - \frac{(\Sigma Y)^2}{n}$$

(Table 2 continued.)

Step # / Source	4	5
Non-simultaneous confidence limits (Procedure C)	$\hat{Y} \pm \sqrt{F_{\gamma,1,n-2}}\,\sqrt{d}\; s_{Y\cdot X}$	
Simultaneous confidence limits	$\hat{Y} \pm \sqrt{2F_{\gamma,2,n-2}}\,\sqrt{d}\; s_{Y\cdot X}$	
Non-simultaneous (P)TL	$\hat{Y} \pm \sqrt{F_{P,1,n-2}}\,\sqrt{1+d}\; s_{Y\cdot X}$	
Simultaneous (P)TL	$\hat{Y} \pm \sqrt{2F_{P,2,n-2}}\,\sqrt{1+d}\; s_{Y\cdot X}$	
Non-simultaneous (γ,P)TL (Procedure B)	$k = \sqrt{F_{P,1,\infty}}\left[1 + \frac{d}{2} - \frac{(2F_{P,1,\infty} - 3)d^2}{24}\right] \cdot \sqrt{F_{\gamma,\infty,n-2}}$	$\hat{Y} \pm k s_{Y\cdot X}$
Simultaneous (γ,P)TL (Procedure D)	$k' = \sqrt{2F_{(1+\gamma)/2,2,n-2}}\,\sqrt{d} + \sqrt{F_{P,1,\infty}F_{(1+\gamma)/2,\infty,n-2}}$	$\hat{Y} \pm k' s_{Y\cdot X}$

Note: F_{λ,ν_1,ν_2} is the λ percentage point of the F distribution with ν_1 and ν_2 degrees of freedom.

It was decided that $P = .95$ and $\gamma = .95$ were reasonable values to use. Figure 3 shows a tolerance band for each of the six types of limits considered in regression when using $P = .95$, $\gamma = .95$ and $n = 15$. Generally all tolerance bands are wide and the price for simultaneity appears high. The cause of the wide limits is two-fold. One cause is that s (basic standard deviation) is perhaps larger than what one would observe under a carefully controlled situation. The second cause of the wide tolerance limits is that either the level of confidence (γ=.95) or the proportion of the population to be included (P=.95) or both were chosen too large in respect to only the 15 pairs of observations used in the sample. In other words, one should pay a high price (large limits) if it is expected that a sample size of 15 should supply the basic information for perhaps hundreds of future predictions.

In order to explore the effect of sample size, it was decided to use the same data under the condition that it were based on 150 pairs of observations rather than only 15 (essentially 10 pairs of observations at each point). Figure 4 shows a band for each of the six types of limits using $P = .95$, $\gamma = .95$ and $n= 150$. From these data one sees a clear distinction between confidence and tolerance bands. The price of simultaneity has become less for both the confidence and the tolerance limits. The non-simultaneous

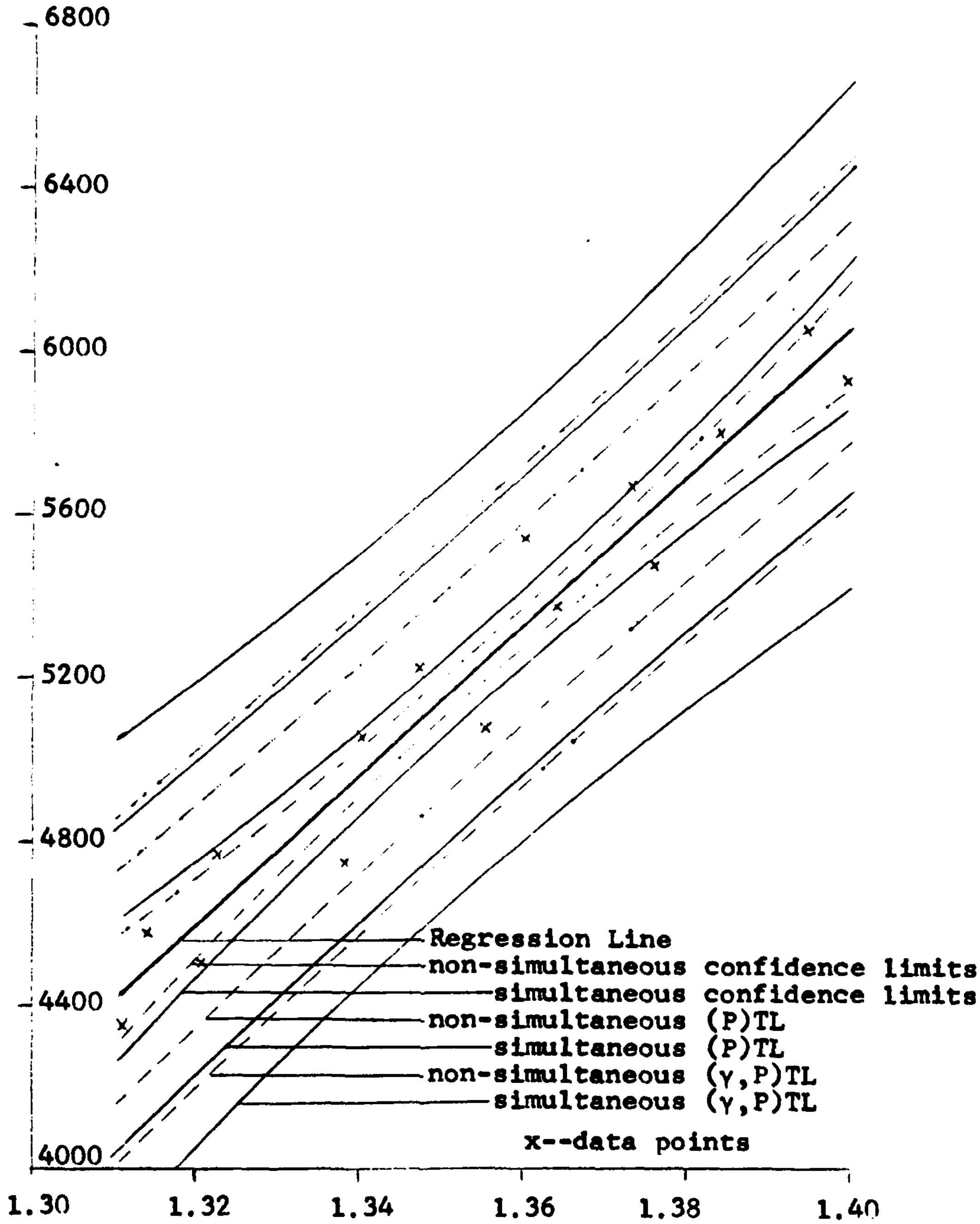

Figure 3. Six Types of Limits for a Simple Linear Regression Problem Using $\gamma=.95$, $P=.95$, and $N=15$.

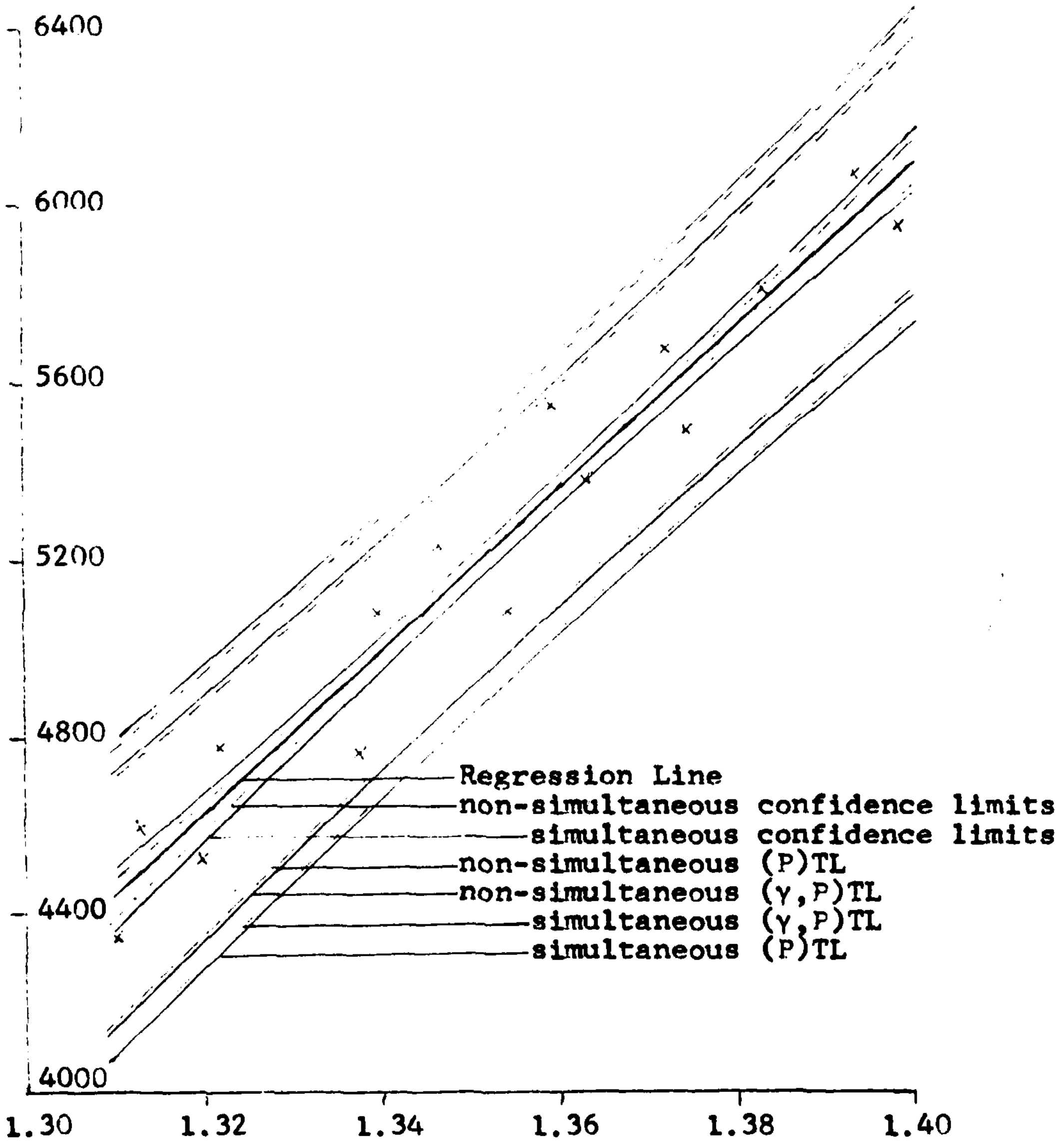

Figure 4. Six Types of Limits for a Simple Linear Regression Problem Using $\gamma=.95$, $P=.95$ and $N=150$ [Essentially 10 pairs/pt.]

(95%) TL do not differ much from the simultaneous (95%)TL. The same is true for the simultaneous and non-simultaneous (95%,95%)TL.

In order to see what role the chosen level of γ plays, it was decided to compute a tolerance band for each of the six types of limits when using $P = .95$, $\gamma = .75$ and $n = 15$. (See Figure 5.) All limits involving γ are about 80% as wide as the limits when using $P=.95$, $\gamma=.95$ and $n= 15$. Of course, both (95%TL) are the same as in Figure 3.

Figure 6 shows the limits for a sample size of 150, $P=.95$ and $\gamma=.75$. Figures 4 and 6 (n=150 for both) are nearly identical. This shows that for a reasonably large sample size the chosen level of γ has very little influence on the width of the confidence or tolerance limits.

Many of the observations made from the sample problem could also be made by comparing the F-ratio values used in the computing formulas in Table 2.

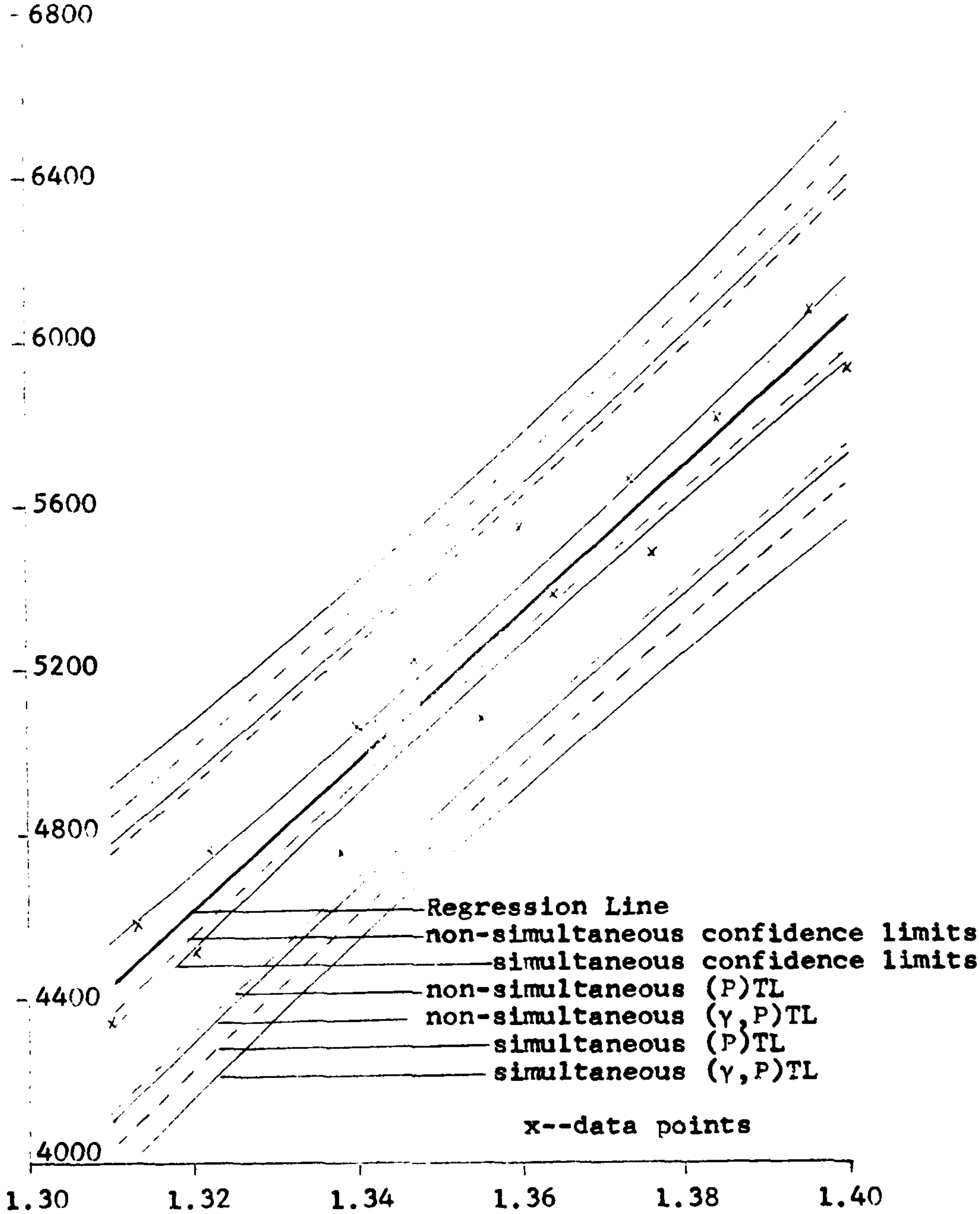

Figure 5. Six Types of Limits for a Simple Linear Regression Problem Using γ=.75, P=.95, and N=15.

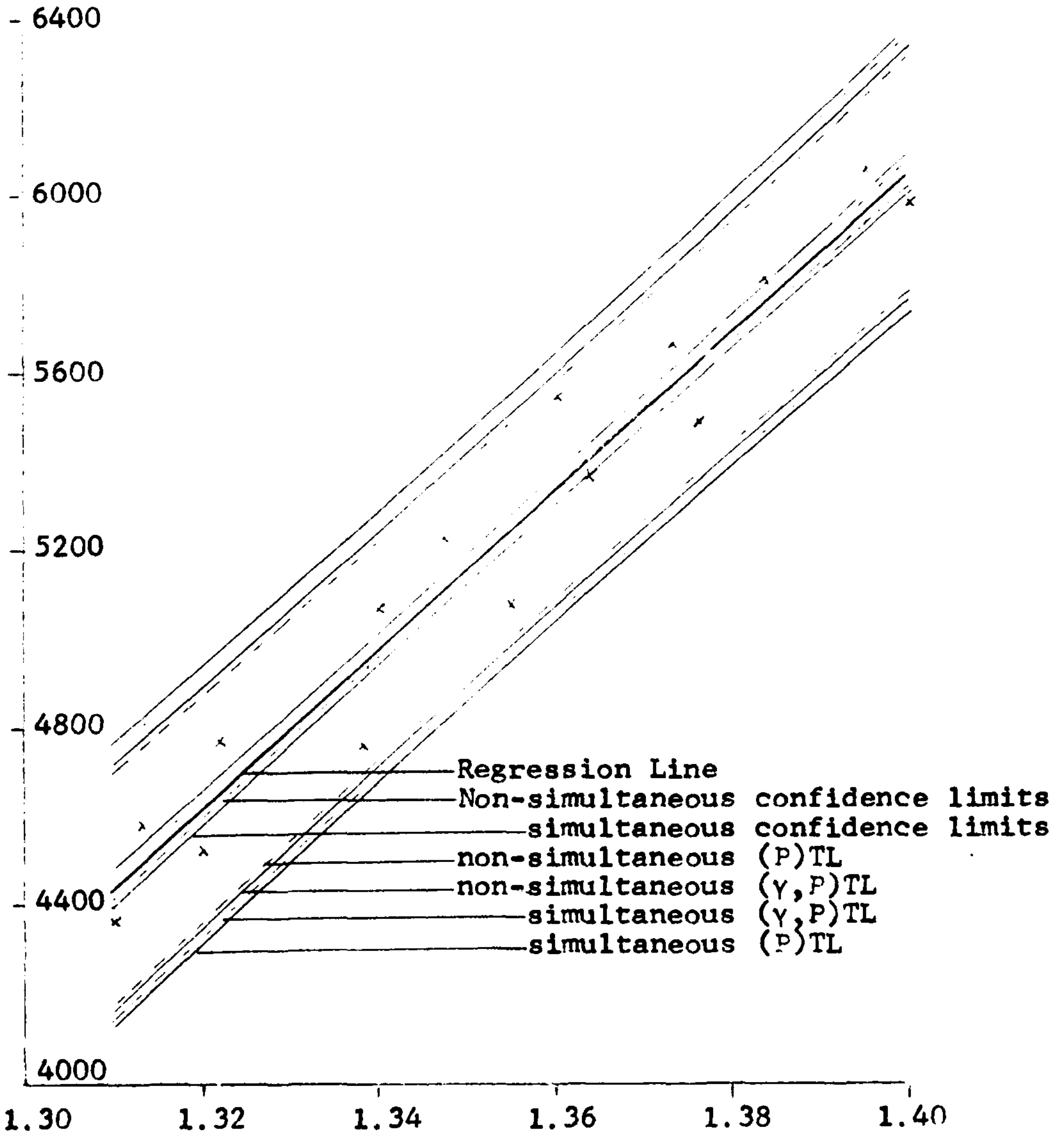

Figure 6. Six Types of Limits for a Simple Linear Regression Problem Using $\gamma=.75$, $P=.95$ and $N=150$ [Essentially 10 pairs/pt.]

VII. RELATED MATERIAL NOT COVERED IN THE PAPER

The material in this paper was limited to two-sided confidence and tolerance limits applied to simple means and simple linear regression lines. Other areas of major interest are:

1. One-sided confidence and tolerance limits.
2. Application of the limits to multiple (fixed X) linear regression problems.
3. Application of the limits to simple linear regression lines where X is measured with error.
4. The simplest of Lieberman & Miller's procedure on simultaneous "P% TL with γ%" was chosen for this paper. Further comparisons between the four procedures under a variety of conditions would be of interest.
5. What price, if any, does the investigator have to pay to be able to make tolerance statements at various values of X <u>not necessarily at the same level P</u>, but still have one over-all γ confidence level compared to a <u>fixed</u> P level statement as given in this report with the same over-all γ level of confidence.
6. Inverse prediction intervals whereby an interval of X values is found for which the additional Y obs. could be associated, and one is 100γ% confident that

at least 100P% of these intervals will include the true associated X_o value (population X_o).

7. Nonparametric confidence and tolerance limits.

VIII. BIBLIOGRAPHY

A. References

1. Acton, F. S. Analysis of straight-line data, pp. 43-50. New York: John Wiley and Sons, Inc., 1959.

2. Bowker, A. H. Computation of factors for tolerance limits on a normal distribution when the sample is large. Ann Math Statist 17:238-240 (1946).

3. Bowker, A. H. Tolerance limits for normal distributions. Chapter 2 of Statistical Research Group, Columbia University, Techniques of Statistical Analysis, pp. 95-110. New York: McGraw-Hill, 1947.

4. Brownlee, K. A. Statistical theory and methodology in science and engineering, pp. 97-99. New York: John Wiley and Sons, Inc., 1960.

5. Brunk, H. D. An introduction to mathematical statistics, pp. 216 and 222. Dallas: Ginn and Company, 1960.

6. Dixon, W. J., and F. J. Massey, Jr. Introduction to statistical analysis, 2nd ed, pp. 195-196. New York: McGraw Hill, 1957.

7. Hald, A. Statistical theory with engineering applications, pp. 311-316. New York: John Wiley and Sons, Inc., 1952.

8. Kendall, M. G., and A. Stuart. The advanced theory of statistics, vol 2, p. 128. New York: Hafner Publishing Co, 1961.

9. Lieberman, G. J. Prediction regions for several predictions from a single regression line. Technometrics 3:21-27 (1961).

10. Lieberman, G. J., and R. G. Miller, Jr. Simultaneous tolerance intervals in regression. Biometrika 50:155-168 (1963).

11. Ostle, B. Statistics in research, pp. 176-177, 325. Ames, Iowa: The Iowa State University Press, 1963.

12. Owen, D. B. Handbook of statistical tables, Section 5.4, 127-137. Reading, Mass.: Addison-Wesley Publishing Co.

13. Paulson, E. A note on tolerance limits. Ann Math Statist 14:90-93 (1943).

14. Proschan, Frank. Confidence and tolerance intervals for the normal distribution. J Amer Statist Assoc 48:550-564 (1953).

15. Scheffé, H. The analysis of variance, pp. 68-70. New York: John Wiley and Sons, Inc., 1959.

16. Steel, R. G. D., and J. H. Torrie. Principles and procedures of statistics, pp. 22 and 73. New York: McGraw-Hill, 1960.

17. Wald, A., and J. Wolfowitz. Tolerance limits for a normal distribution. Ann Math Statist 17:208-215 (1946).

18. Wallis, W. Allen. Tolerance intervals for linear regression. Second Berkeley Symposium on Mathematical Statistics and Probability, edited by Jerzy Neyman. Berkeley: University of California Press, pp. 43-51 (1951).

19. Weissberg, A., and G. H. Beatty. Tables of tolerance-limit factors for normal distributions. Technometrics 2:483-500 (1960).

20. Wilks, S. S. Determination of sample sizes for setting tolerance limits. Ann Math Statist 12:91-96 (1941).

21. Wilks, S. S. Mathematical statistics, pp. 291-294. New York: John Wiley and Sons, Inc., 1962.

22. Working, H., and H. Hotellings. Applications of the theory of error to the interpretation of trends. J Amer Statist Assoc 24:73-85 (1929).

B. Additional Bibliography

The following articles, although not cited specifically in this thesis, discuss additional topics on confidence and tolerance limits (regions).

1. Parametric Confidence Limits

Banerjee, S. K. Approximate confidence interval for linear functions of means of k populations when the populations variances are not equal. Sankhyā 22:357-358 (1960).

Banerjee, S. K. Expressions for the lower bound to confidence co-efficients. Sankhyā 21:127-140 (1959).

Banerjee, S. K. On confidence interval for the two-mean problem based on separate estimates of variances and tabulated values of t-table. Sankhyā, A 23:359-379 (1961).

Bartlett, M. S. Approximate confidence intervals. Biometrika 40:12-19 (1953).

Bartlett, M. S. Approximate confidence intervals. II. More than one unknown parameter. Biometrika 40:306-317 (1953).

Bartlett, M. S. Approximate confidence intervals. III. A bias correction. Biometrika 42:201-204 (1955).

Bennett, B. M. On the performance characteristic of certain methods of determining confidence limits. Sankhyā 18:1-12 (1957).

Brillinger, D. R. The asymptotic behavior of Tukey's general method of setting approximate confidence limits (the jackknife) when applied to maximum likelihood estimates. Rev Inst Internat Statist 32:202-206 (1964).

Bunke, O. New confidence intervals for the parameters of the binomial distribution. Wiss Zeit Humboldt Univ, Math-Nat Reihe 9:335-363 (1959).

Correa Po'lit, H. Statistical inference about the parameters of nonnormal populations (confidence intervals). Trabajos Estadist 9:118-140 (1958).

Creasy, M. A. Confidence limits for the gradient in the linear functional relationship. J Roy Stat Soc B 18:65-69 (1956).

Dorogovcev, A. Ya. Confidence intervals in estimation of parameters. Dopovidi Akad Nauk Ukrain RSR pp. 355-358, 1959.

Dubey, S. On the determination of confidence limits of an index. Biometrics 22:603-609 (1966).

Dunn, O. J. Confidence intervals for the means of dependent, normally distributed random variables. J Amer Statist Assoc 54:613-621 (1959).

Farrell, R. H. Sequentially determined bounded length confidence intervals. Ph.D. thesis, 1959, University of Illinois.

Farrell, R. H. Bounded length confidence intervals for the zero of a regression function. Ann Math Statist 33:237-247 (1962).

Farrell, R. H. Bounded length confidence intervals for the p-point of a distribution function, II. Ann Math Statist 37:581-585 (1966).

Farrell, R. H. Bounded length confidence intervals for the p-point of a distribution function, III. Ann Math Statist 37:586-592 (1966).

Goldman, A. Sample size for a specified width confidence interval on the ratio of variances from two independent normal populations. Biometrics 19:465-477 (1963).

Guenther, W. C., and M. G. Whitcomb. Critical regions for tests of interval hypotheses about the variance. J Amer Statist Assoc 61:204-219 (1966).

Halperin, M. Confidence interval estimation in non-linear regression. J Roy Stat Soc B 25:330-333 (1963).

Halperin, M. Confidence intervals from censored samples. Ann Math Statist 32:828-837 (1961).

Halperin, M. Interval estimation on non-linear parametric functions. J Amer Statist Assoc 59:168-181 (1964).

Halperin, M. Interval estimation of non-linear parametric functions, III. J Amer Statist Assoc 60:1191-1199 (1965).

Halperin, M. Note on interval estimation in non-linear regression when responses are correlated. J Roy Stat Soc B 26:267-269 (1964).

Hamaker, H. C. Average confidence limits for binomial probabilities. Rev Inst Internat Statistique 21:17-27 (1953).

Harter, H. L. Criteria for best interval estimators. Bull Int Statist Inst 40:766 (1964).

Huitson, A. A method of assigning confidence limits to linear combinations of variances. Biometrika 42:471-479 (1955).

Huzurbazar, V. S. Confidence intervals for the parameter of a distribution admitting a sufficient statistic when the range depends on the parameter. J Roy Stat Soc B 17:86-90 (1955).

Koopmans, L. H., D. B. Owen, and J. I. Rosenblatt. Confidence intervals for the coefficients of variation for the normal and log normal distributions. Biometrika 51:25-32 (1964).

Kraemer, H. C. One-sided confidence intervals for the quality indices of a complex item. Technometrics 5:400-403 (1963).

Kramer, K. H. Tables for constructing confidence limits on the multiple correlation coefficient. J Amer Statist Assoc 58:1082-1085 (1963).

Linhart, H. Approximate confidence limits for the coefficient of variation of gamma distributions. Biometrics 21:733-738 (1965).

Madansky, A. More on length of confidence intervals. J Amer Statist Assoc 57:586-589 (1962).

McHugh, R. B. Confidence interval inference and sample size determination. American Statistician 15:14-17 (1961).

Moriguti, S. Confidence limits for a variance component. Rep Statist Appl Res Union Jap Sci Eng 3:29-41 (1954).

Natrella, M. G. The relationship between confidence intervals and tests of significance - a teaching aid. American Statistician 14:20-22 (1960).

Ogawa, J. On a confidence interval of the ratio of population means of a bivariate normal distribution. Proc Japan Acad 27:313-316 (1951).

Peers, H. W. On confidence points and Bayesian probability points in the case of several parameters. J Roy Stat Soc B 27:9-16 (1965).

Pillai, K. C. S. Confidence interval for the correlation coefficient. Sankhyā 7:415-422 (1946).

Pratt, J. W. Length of confidence intervals. J Amer Statist Assoc 56:549-567 (1961).

Pratt, J. W. Shorter confidence intervals for the mean of a normal distribution with known variance. Ann Math Statist 34:574-586 (1963).

Press, S. J. A confidence interval comparison of two test procedures for the Behrens-Fisher Problem. J Amer Statist Assoc 61:454-466 (1966).

Ray, W. D. Sequential confidence intervals for the mean of a normal population with unknown variance. J Roy Stat Soc B 19:133-143 (1957).

Sandelius, M. A confidence interval for the smallest proportion of a binomial population. J Roy Stat Soc B 14:115-116 (1952).

Scheffé, H. A method for judging all contrasts in the analysis of variance. Biometrika 40:87-104 (1953).

Scheffé, H. Note on the use of the tables of percentage points of the incomplete beta function to calculate small sample confidence intervals for a binomial p. Biometrika 33:181 (1944).

Simonds, T. A. Mean time between failure (MTBF) confidence limits. Indust Qual Contr 20:21-27 (1963).

Siotani, M. Interval estimation for linear combinations of means. J Amer Statist Assoc 59:1141-1164 (1964).

Stevens, W. L. Shorter intervals for the parameter of the binomial and Poisson distributions. Biometrika 44:436-440 (1957).

Tate, R. F., and G. W. Klett. Optimal confidence intervals for the variance of a normal distribution. J Amer Statist Assoc 54:674-682 (1959).

Terpstra, T. J. A confidence interval for the probability that a normally distributed variable exceeds a given value, based on the mean and the mean range of a number of samples. Appl Sci Research A 3:297-307 (1952).

Thatcher, A. R. Relationship between Bayesian and confidence limits for predictions. J Roy Stat Soc B 26:176-192 (1964).

Tukey, J. W., and D. H. McLaughlin. Less vulnerable confidence and significance procedures for location based on a single sample: Trimming/Winsorisation. Sankhyā A 25:331-352 (1963).

Welch, B. L. On comparisons between confidence point procedures in the case of a single parameter. J Roy Stat Soc B 27:1-8 (1965).

Williams, J. S. A confidence interval for variance components. Biometrika 49:278-281 (1962).

Wuler, H. Confidence limits for the mean of a normal population with known coefficient of variation. Aust J Applied Sci 9:321-325 (1958).

2. *Non-parametric Confidence Limits*

Bennett, B. M. Confidence limits for a ratio using Wilcoxon's signed rank test. Biometrics 21:231-234 (1965).

Harter, H. L. Exact confidence bounds, based on one order statistic, for the parameter of an exponential population. Technometrics 6:301-317 (1964).

Lehmann, E. L. Nonparametric confidence intervals for a shift parameter. Ann Math Statist 34:1507-1512 (1963).

Leone, F. C., Y. H. Rutenberg, and C. W. Topp. The use of sample quasi-ranges in setting confidence intervals for the population standard deviation. J Amer Statist Assoc 56:260-272 (1961).

McCarthy, P. J. Stratified sampling and distribution-free confidence intervals for a median. J Amer Statist Assoc 60:772-783 (1965).

Nair, K. R. Table of confidence interval for the median in samples from any continuous population. Sankhyā 4:551-558 (1940).

Noether, G. E. Wilcoxon confidence intervals for location parameters in the discrete case. J Amer Statist Assoc 62:184-188 (1967).

Rosenblatt, J. Tests and confidence intervals based on the metric d_2. Ann Math Statist 34:618-623 (1963).

Walsh, J. E. Distribution-free tolerance intervals for continuous symmetrical populations. Ann Math Statist 33:1167-1174 (1962).

Walsh, J. E. Large sample confidence intervals for density function values at percentage points. Sankhyā 12:265-276 (1953).

Weiss, L. Confidence intervals of preassigned length for quantiles of unimodal populations. Naval Res Log Quart 7:251-256 (1960).

Welch, B. L., and H. W. Peers. On formulae for confidence points based on integrals of weighted likelihood. J Roy Stat Soc B 25:318-329 (1963).

Wilson, E. B. On confidence intervals. Proc Nat Acad Sci USA 28:88-93 (1942).

3. Confidence Regions (Simultaneous Confidence Limits)

Aitchison, J. Confidence-region tests. J Roy Stat Soc B 26:462-476 (1964).

Aitchison, J. Likelihood ratio and confidence-region tests. J Roy Stat Soc B 27:245-250 (1965).

Beale, E. M. L. Confidence regions in non-linear estimation. J Roy Stat Soc B 22:41-76, 77-78 (1960).

Borges, R. Subjective most selective confidence regions. Zeit Wahrscheinlichkeits 1:47-69 (1962).

Bowden, D. C., and F. A. Graybill. Confidence bands of uniform and proportional width for linear models. J Amer Statist Assoc 61:182-198 (1966).

Box, G. E. P., and J. S. Hunter. A confidence region for the solution of a set of simultaneous equations with an application to experimental design. Biometrika 41:190-199 (1954).

Chatterjee, S. K. Simultaneous confidence intervals of predetermined length based on sequential samples. Bull Calcutta Statist Assoc 11:144-149 (1962).

Chew, V. Confidence, prediction, and tolerance regions for the multivariate normal distribution. J Amer Statist Assoc 61:605-617 (1966).

Dwass, M. Multiple confidence procedures. Ann Inst Statist Math, Tokyo, 10:277-282 (1959).

Dunn, O. J. Multiple comparison among means. J Amer Statist Assoc 56:52-64 (1961).

Erlander, S., and J. Gustavsson. Simultaneous confidence regions in normal regression analysis with an application to road accidents. Rev Inst Internat Statist 33:364-377 (1965).

Gafarian, A. V. Confidence bands in straight line regression. J Amer Statist Assoc 59:182-213 (1964).

Goodman, L. A. On simultaneous confidence intervals for multinomial proportions. Technometrics 7:247-254 (1965).

Goodman, L. A. Simultaneous confidence intervals for contrasts among multinomial populations. Ann Math Statist 35:716-725 (1964).

Goodman, L. A. Simultaneous confidence limits for cross-product ratios in contingency tables. J Roy Stat Soc B 26:86-102 (1964).

Hartley, H. O. Exact confidence regions for the parameters in non-linear regression laws. Biometrika 51:347-354 (1964).

Hemelrijk, J. Construction of a confidence region for a line. Nederl Akad Wetensch, Proc 52:374-384, 995-1005, Indagationes Math 11 (1949).

Hoel, P. G. Confidence regions for linear regression. Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950, pp. 75-81. Berkeley and Los Angeles: University of California Press, 1951.

Khatri, C. G. Simultaneous confidence bounds connected with a general linear hypothesis. J Maharaja Sayajirao Univ Baroda 10:11-13, No. 3 (1962).

Kunisawa, Kiyonori, Makabe, Hajime, and Morimura, Hidenori. Tables of confidence bands for the population distribution function I. Rep Statist Appl Res Union Jap Sci Eng 1:23-44 (1951).

Maritz, J. S. Confidence regions for regression parameters. Aust J Statist 4:4-10 (1962).

Quesenberry, C. P., and D. C. Hurst. Large sample simultaneous confidence intervals for multinomial proportions. Technometrics 6:191-195 (1964).

Ramachandran, K. V. Contributions to simultaneous confidence interval estimation. Biometrics 12:51-56 (1956).

Roy, S. N. A survey of some recent results in normal multivariate confidence bounds. Bull Int Statist Inst 39 II:405-422 (1962).

Roy, S. N., and R. C. Bose. Simultaneous confidence interval estimation. Ann Math Statist 24:513-536 (1953).

Roy, S. N., and R. Gnanadesikan. A note on further contributions to multivariate confidence bounds. Biometrika 45:581 (1958).

Roy, S. N., and R. Gnanadesikan. Further contributions to multivariate confidence bounds. Biometrika 44:289-292 (1957).

Scheffé, H. Simultaneous interval estimate of linear functions of parameters. Bull Int Statist Inst 38 IV:245-253 (1961).

Sen, P. K. On nonparametric simultaneous confidence regions and tests for the one criterion analysis of variance problem. Ann Inst Statist Math, Tokyo 18:319-335 (1966).

Siotani, Minoru, Kawakami, and Hisoka. Simultaneous confidence interval estimation on regression coefficients. Proc Inst Statist Math 10:79-98 (1962).

Smirnov, N. V. On the construction of confidence regions for the density of distribution of random variables. Doklady Akad Nauk SSSR (N.S.) 74:189-191 (1950) Russian.

Stein, C. M. Confidence sets for the mean of a multivariate normal distribution. J Roy Stat Soc B 24:265-296 (1962).

Tukey, J. W. The problem of multiple comparisons. MS of 396 pages, Princeton Univ (1953).

Verma, M. C., and M. N. Ghosh. Simultaneous tests of linear hypothesis and confidence interval estimation. J Indian Soc Agric Statist 15:194-211 (1963).

Walsh, J. E. Simultaneous confidence intervals for differences of classification probabilities. Biom Zeit 5:231-234 (1963).

4. Parametric Tolerance Limits and Regions

Aitchison, J. Bayesian tolerance regions. J Roy Stat Soc B 26:161-175 (1964).

Aitchison, J. Expected-cover and linear-utility tolerance intervals. J Roy Stat Soc 28:57-62 (1966).

Bain, L. J., and D. L. Weeks. Tolerance limits for the generalized gamma distribution. J Amer Statist Assoc 60:1142-1152 (1965).

Barlow, R. E., and F. Proschan. Tolerance and confidence limits for classes of distributions based on failure rate. Ann Math Statist 37:1593-1601 (1966).

Ellison, B. E. On two-sided tolerance intervals for a normal distribution. Ann Math Statist 35:762-772 (1964).

Epstein, B. Tolerance limits based on life test data taken from an exponential distribution. Indust Qual Cont 17:10-11 (1960).

Guttmann, I. Best population and tolerance regions. Ann Inst Statist Math 13:9-26 (1961).

Hanson, D. L., and L. H. Koopmans. Tolerance limits for the class of distributions with increasing hazard rates. Ann Math Statist 35:1561-1570 (1964).

Ireson, W. G., G. J. Resnikoff, and B. E. Smith. Statistical tolerance limits for determining process capability. J Industr Engineering 12:126-131 (1961).

Iskii, Goro, Kudo, Hirokichi. Tolerance region for missing variables in linear statistical models. J Math Osaka City Univ 14:117-130 (1963).

Jilek, M., and O. Likar. Statistical tolerances. Wis Zeit Techn, Univ Dresden 11:1253-1256 (1962).

Jilek, M., and O. Likar. Tolerance limits of the normal distribution with known variance and unknown mean. Aust J Statist 2:78-83 (1960).

Jilek, M., and O. Likar. Tolerance regions of the normal distribution with known μ and unknown σ. Biom Zeit 2:204-209 (1960).

John, S. A tolerance region for multivariate normal distribution. Sankhyā A 25:363-368 (1963).

Mitra, S. K. Tables for tolerance limits for a normal population based on sample mean and range or mean range. J Amer Statist Assoc 52:88-94 (1957).

Mouradian, G. Tolerance limits for assemblies and engineering relationships, pp. 598-606. Annual Tech Conf Trans, Am Soc Quality Control, New York, N.Y. (1966).

Seeger, P. Some examples involving the setting of tolerance limits. Nordisk Tidskrift för Industriel Statistik 10:1-7 (1966).

Wolfowitz, J. Confidence limits for the fraction of a normal population which lie between two given limits. Ann Math Statist 17:483-488 (1946).

5. Non-parametric Tolerance Limits and Regions

Goodman, L., and A. Madansky. Parameter-free and non-parametric tolerance limits: the exponential case. Technometrics 4:75-95 (1962).

Hanson, D. L., and D. B. Owens. Distribution-free tolerance limits elimination of the requirement that cumulative distribution be continuous. Technometrics 5:518-521 (1963).

Jivina, M. Sequential estimation of distribution-free tolerance limits. Cehoslovack Mat Z 2:221, 231 (77) (1952); correction 3:283 (78) (1953).

Nelson, L. S. Nomograph for two-sided distribution-free tolerance intervals. Industr Qual Contr XII 19:11-13 (1963).

Walsh, J. E. Statistical prediction from tolerance regions. Bull Int Statist Inst III 38:313-317 (1961).

Walsh, J. E. Some two-sided distribution-free tolerance intervals of a general nature. J Amer Statist Assoc 57:775-784 (1962).

6. Books

Many books have a section on confidence limits and some have a section on tolerance limits. The books listed have either a considerable amount of information on confidence and tolerance limits or they cover material not covered in the given articles.

Brownlee, K. A. Statistical theory and methodology in science and engineering, pp. 284-288. New York: John Wiley and Sons, Inc., 1960.

Chung, J. H., and D. B. DeLury. Confidence limits for the hypergeometric distribution, University of Toronto Press, 1950.

Diem, Konrad (Editor). Scientific tables, Geigy Pharmaceuticals, Ardlsey, New York, 6th Edition.

Finney, D. J. Statistical methods in biological assay. New York: Hafner Publishing Co., 1952.

Owen, D. B. Factors for one-sided tolerance limits and for variables sampling plans. Sandia Corporation Monograph SCR-607 (1963).

Wilks, S. S. Mathematical statistics. New York: John Wiley and Sons, Inc., 1962.

7. Indexes to Journals

None of the articles given in the indexes are referred to specifically in the additional bibliography.

Clunies-Ross, C. W., W. A. Glenn, and L. S. Brenna. Index to the Journal of the American Statistical Association, Volumes 35-50 (1940-1955), Confidence pp. 52-53 (44 references), Tolerance p. 129 (4 references).

Greenwood, J. A., I. Olkin, and I. R. Savage. The Annals of Mathematical Statistics, Indexes to Volumes 1-31 1930-1960. Confidence intervals and regions, pp. 293-299 (368 references), Tolerance limits and regions, pp. 523-524 (91 references).

Matusita, K. Annals of the Institute of Statistical Mathematics--Contents Vol 11 (1959) - Vol 17 (1965), Estimation pp. 36-38 (10 references to confidence and tolerance limits (regions).

Merrington, Maxine, Biometrika Index, 1955, Volumes 1-37 (1901-1950) confidence intervals (13 references).

DATE
FILMED
8-8